

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 21:06:50 ; Search time 175.499 Seconds  
(without alignments)  
8457.264 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_5\_55  
Perfect score: 51  
Sequence: 1 ttaagaccagcctggacaac.....sggtctctacaaaaataca 51  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_da.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_lmd.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pi.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vl.\*
- 30: em\_hgt\_hum.\*
- 31: em\_hgt\_inv.\*
- 32: em\_hgt\_other.\*
- 33: em\_hgt\_mus.\*
- 34: em\_hgt\_pln.\*
- 35: em\_hgt\_rtd.\*
- 36: em\_hgt\_mam.\*
- 37: em\_hgt\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgtgo\_hum.\*
- 40: em\_hgtgo\_mus.\*
- 41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50.6	99.2	1669	6	AX207224 Sequence
2	50.6	99.2	1680	6	AX394457 Sequence
3	50.6	99.2	9432	6	AX394456 Sequence
4	50.6	99.2	9432	9	HUMCYP2D6
5	50.6	99.2	13677	9	HSCYP2D7B
6	44.6	87.5	173728	9	AL365496 Human DNA
7	44.2	86.7	38740	9	HSU59962 Human cosmi
8	44.2	86.7	69185	2	AC022117 Homo sapi
9	44.2	86.7	86424	2	AC010222 Homo sapi
10	44.2	86.7	97559	9	AL356748 Human DNA
11	44.2	86.7	97860	9	AC091816 Homo sapi
12	44.2	86.7	101076	9	HS37M17
13	44.2	86.7	106314	9	AC015669 Homo sapi
14	44.2	86.7	124328	9	AL133461 Human DNA
15	44.2	86.7	130965	9	AC026773 Homo sapi
16	44.2	86.7	139714	2	AC090364 Homo sapi
17	44.2	86.7	142305	9	AC013476 Homo sapi
18	44.2	86.7	148043	9	AC008541 Homo sapi
19	44.2	86.7	166288	9	AL139410 Human DNA
20	44.2	86.7	167746	9	AC011120 Homo sapi
21	44.2	86.7	167758	2	AC068746 Homo sapi
22	44.2	86.7	168069	9	AC093520 Homo sapi
23	44.2	86.7	168672	2	AC024920 Homo sapi
24	44.2	86.7	171112	9	AC016866 Homo sapi
25	44.2	86.7	178418	2	AC013243 Homo sapi
26	44.2	86.7	183085	9	AC005815 Homo sapi
27	44.2	86.7	186299	2	AC107202 Homo sapi
28	44.2	86.7	187229	9	AC027672 Homo sapi
29	44.2	86.7	189947	2	AC048338 Homo sapi
30	44.2	86.7	194143	2	AC022222 Homo sapi
31	44.2	86.7	214445	2	AL365497 Homo sapi
32	44.2	86.7	219554	2	AC107016 Homo sapi
33	43.2	84.7	186084	9	AP002768 Homo sapi
34	43.2	84.3	147594	9	AC013829 Homo sapi
35	43.2	84.3	183072	2	AC073594 Homo sapi
36	43.2	84.3	184000	2	AC010892 Homo sapi
37	43.2	84.3	185029	2	AC068986 Homo sapi
38	43.2	84.3	203038	9	AC076968 Homo sapi
39	42.8	83.9	141867	2	AC006209 Homo sapi
40	42.6	83.5	401.6	6	AX270321 Sequence
41	42.6	83.5	401.6	6	AX271852 Sequence
42	42.6	83.5	565	6	AX388115 Sequence
43	42.6	83.5	2646	9	AK096230 Homo sapi
44	42.6	83.5	4278	9	HSMB04155 Homo sapi
45	42.6	83.5	12047	6	AR078691 Sequence

ALIGNMENTS

RESULT 1	AX207224	AX207224	Sequence 1	1669 bp	DNA	linear	PAT 30-AUG-2001
AX207224	LOCUS	Sequence 1	from Patent WO0155432.				
	DEFINITION	AX207224					
	ACCESSION	AX207224					
	VERSION	AX207224.1	GI:15394976				
	KEYWORDS						
	SOURCE						
	ORGANISM						
	REFERENCE						
	AUTHORS						
	TITLE						
	JOURNAL						

AX207224  
Sequence 1  
from Patent WO0155432.  
AX207224  
AX207224.1  
GI:15394976  
synthetic construct.  
synthetic construct  
artificial sequences.  
1 (bases 1 to 1669)  
Raimundo, S. and Zanger, U.  
Polymorphisms in the human cyp2d6 gene promoter region and their  
use in diagnostic and therapeutic applications  
Patent: WO 0155432-A 1 02-AUG-2001;

```

FEATURES
  source
    Epidauros Biotechnologie AG (DE)
    Location/Qualifiers
      1..1669
      /organism="synthetic construct"
      /db_xref="taxon:32630"
      /note="artificial sequence"
BASE COUNT
  413 a 376 c 534 g 338 t      8 others
ORIGIN
  Query Match
    99.2%; Score 50.6; DB 6; Length 1669;
  Best Local Similarity
    100.0%; Pred. No. 3.9e-07;
  Matches
    51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGACAACTTGAAGAACCCGGTCTCTACAAAAATACA 51
    |||||
Db 5 TCAAGACCGCTGGACAACTTGAAGAACCCGGTCTCTACAAAAATACA 55

RESULT 2
AX394457
LOCUS
  AX394457
DEFINITION
  Sequence 2 from Patent WO0218638.
ACCESSION
  AX394457
VERSION
  AX394457.1 GI:21065595
KEYWORDS
  human.
ORGANISM
  Homo sapiens
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1
AUTHORS
  Risinger,C., Andersson,M.K., Lewander,T. and Ollasson,E.
TITLE
  Detection of cyp2d6 polymorphisms
JOURNAL
  Patent: WO 0218638-A 2 07-MAR-2002;
  Gemini Genomics PLC (GB)
FEATURES
  source
    1..1680
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT
  413 a 379 c 539 g 342 t      7 others
ORIGIN
  Query Match
    99.2%; Score 50.6; DB 6; Length 1680;
  Best Local Similarity
    100.0%; Pred. No. 3.9e-07;
  Matches
    51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGACAACTTGAAGAACCCGGTCTCTACAAAAATACA 51
    |||||
Db 5 TCAAGACCGCTGGACAACTTGAAGAACCCGGTCTCTACAAAAATACA 55

RESULT 3
AX394456
LOCUS
  AX394456
DEFINITION
  Sequence 1 from Patent WO0218638.
ACCESSION
  AX394456
VERSION
  AX394456.1 GI:21065594
KEYWORDS
  human.
ORGANISM
  Homo sapiens
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1
AUTHORS
  Risinger,C., Andersson,M.K., Lewander,T. and Ollasson,E.
TITLE
  Detection of cyp2d6 polymorphisms
JOURNAL
  Patent: WO 0218638-A 1 07-MAR-2002;
  Gemini Genomics PLC (GB)
FEATURES
  source
    1..9432
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
BASE COUNT
  1964 a 2647 c 2976 g 1845 t
ORIGIN

```

```

Query Match
  99.2%; Score 50.6; DB 6; Length 9432;
Best Local Similarity
  98.0%; Pred. No. 3.4e-07;
Matches
  50; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGACAACTTGAAGAACCCGGTCTCTACAAAAATACA 51
    |||||
Db 5 TCAAGACCGCTGGACAACTTGAAGAACCCGGTCTCTACAAAAATACA 55

RESULT 4
HUMCYP2D6
LOCUS
  HUMCYP2D6
DEFINITION
  Human cytochrome P450 IID6 (CYP2D6) gene, complete cds.
ACCESSION
  M33388
VERSION
  M33388.1 GI:181303
KEYWORDS
  cytochrome P450; cytochrome P450 IID6.
SOURCE
  Human DNA, clone lambda2D-18/2.
ORGANISM
  Homo sapiens
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 9432)
AUTHORS
  Kimura,S., Umeno,M., Skoda,R.C., Meyer,U.A. and Gonzalez,F.J.
TITLE
  The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and
  identification of the polymorphic CYP2D6 gene, a related gene, and
  a pseudogene
JOURNAL
  Am. J. Hum. Genet. 45 (6), 889-904 (1989)
MEDLINE
  90072069
PUBMED
  2574001
COMMENT
  Draft entry and computer-readable sequence for [Am. J. Hum. Genet.
  45, 889-904 (1989)] kindly submitted
  by S.Kimura, 29-MAR-1990.
FEATURES
  Location/Qualifiers
    1..9432
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /map="22q13.1"
    join(1532..1799,2503..2674,3225..3377,3466..3626,
    4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)
    /gene="CYP2D6"
    join(1532..1799,2503..2674,3225..3377,3466..3626,
    4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)
    /product="cytochrome P450 IID6"
    /note="G00-132-127"
    1532..1799
    /partial
    /gene="CYP2D6"
    /note="cytochrome P450 IID6; G00-132-127"
    /number=1
    join(1620..1799,2503..2674,3225..3377,3466..3626,
    4060..4236,4427..4568,4776..4963,5418..5559,5658..5836)
    /gene="CYP2D6"
    /codon_start=1
    /product="cytochrome P450 IID6"
    /protein_id="AA53500.1"
    /db_xref="GI:181304"
    /db_xref="GDB:G00-132-127"
    /translation="MGLEALVPLAVIPIFLLLVLMHRRQRWAARYPPGPLPLGLG
    NLHVDPTNPCTDQLRRFGDVSLOQAWTPVVLNGLAAVREALVTHGEDTADRP
    FVPTQILGFGRPSQGVFLYRQPAWREQRFSVTLNLGLGKSLQEWVTEAACL
    CAAFANHGSRPFRNGLLDKAVSNVTSLTGGRREYDDPRFLRLDLAQEGIKESG
    FLREVLNAVPLLIIPALAGKLVRFQKAFITOLDLLETHRTWDPQPRDLTEAFL
    AEMEKAGNPESSENENLRIVVADLFSGMVTTSITLAWLLMLIHFDVQRVQOE
    IDDVIGVREPENGQAHMPYTTAVIHEVQFGDIVPLGVTHMTSRDIEVQGFRIKPG
    TTLITLSSVLKDEAVWEKFRFHPHFHDAQGHFVKPEAFLPFSAGRACLGEPLAR
    TELTFLFTSLLLQHFSSVTPGQPRPHSHGVFAFLVSPSPYELCAVRP"
    1800..2502
    /gene="CYP2D6"
    /note="G00-132-127"
    /number=1
    2503..2674
    /gene="CYP2D6"
    /note="G00-132-127"
  intron
  exon

```



intron /number=2  
2675..3224  
/gene="CYP2D6"  
/note="G00-132-127; does not fit consensus"  
exon /number=2  
3225..3377  
/gene="CYP2D6"  
/note="G00-132-127"  
intron /number=3  
3378..3465  
/gene="CYP2D6"  
/note="G00-132-127"  
exon /number=3  
3466..3626  
/gene="CYP2D6"  
/note="G00-132-127"  
intron /number=4  
3627..4059  
/gene="CYP2D6"  
/note="G00-132-127"  
exon /number=4  
4060..4236  
/gene="CYP2D6"  
/note="G00-132-127"  
intron /number=5  
4237..4426  
/gene="CYP2D6"  
/note="G00-132-127"  
exon /number=5  
4427..4568  
/gene="CYP2D6"  
/note="G00-132-127"  
intron /number=6  
4569..4775  
/gene="CYP2D6"  
/note="G00-132-127"  
exon /number=6  
4776..4963  
/gene="CYP2D6"  
/note="G00-132-127"  
intron /number=7  
4964..5417  
/gene="CYP2D6"  
/note="G00-132-127"  
exon /number=7  
5418..5559  
/gene="CYP2D6"  
/note="G00-132-127"  
intron /number=8  
5560..5657  
/gene="CYP2D6"  
/note="G00-132-127"  
exon /number=8  
5658..5909  
/partial  
/gene="CYP2D6"  
/note="G00-132-127"  
BASE COUNT 1964 a 2647 c 2976 g 1845 t  
ORIGIN

DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Human CYP2D7BP pseudogene for cytochrome P450 2D6.  
X58468  
X58468.1 GI:30337  
CYP2D7BP gene; Cytochrome P450; cytochrome P450 2D6; pseudogene.  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Heim, M.H.  
1 (bases 1 to 13677)  
Direct Submission  
Submitted (25-MAR-1991) M.H. Helm, Dept of Pharmacology, Biocentre  
University of Basel, Klingelbergstr 70, 4056 Basel, SWITZERLAND  
Heim, M.H. and Meyer, U.A.  
2 (bases 1 to 13677)  
Evolution of a highly polymorphic human cytochrome P450 gene  
Cluster: CYP2D6  
Genomics 14 (1), 49-58 (1992)  
93052308  
1358797  
See X58467, and Am. J. Hum. Genet. 47:994-1001(1990).

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REFERENCE  
AUTHORS  
TITLE

JOURNAL  
MEDLINE  
PUBMED

COMMENT  
FEATURES

Source

1. .13677  
Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="22"

/clone="45"

/clone\_lib="EMBL4"

/dev\_stage="adult"

1534..5868

/gene="CYP2D7BP"

join(1534..1801,2504..2675,3203..3355,3444..3605,

4031..4207,4400..4540,4735..4922,5377..5518,5617..5868)

gene

mRNA

exon

CDS

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

intron

exon

RESULT 5  
HSCYP2D7B  
LOCUS HSCYP2D7B 13677 bp DNA linear PRI 21-OCT-1992

Query Match 99.2%; Score 50.6; DB 9; Length 9432;  
Best Local Similarity 98.0%; Pred. No. 3 4e-07;  
Matches 50; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCAAGACCGCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 5 TCAAGACCGCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 55  
|||||



```
repeat_region 9780..10168
repeat_region /rpt_family="THB1b"
repeat_region complement(11057..11143)
repeat_region /rpt_family="MER1"
repeat_region complement(11157..11388)
repeat_region /rpt_family="MER1"
STS 11378..11615
/standard_name="AFW211wb6"
repeat_region /note="D12598"
13016..13298
repeat_region /rpt_family="ALU"
repeat_region complement(14487..14781)
repeat_region /rpt_family="ALU"
repeat_region complement(15614..15893)
repeat_region /rpt_family="ALU"
repeat_region complement(16162..16280)
repeat_region /rpt_family="MST"
repeat_region complement(17272..17646)
repeat_region /rpt_family="MER42"
repeat_region complement(19390..20690)
repeat_region /rpt_family="L1"
20691..22327
repeat_region /rpt_family="L1"
repeat_region complement(24243..24516)
repeat_region /rpt_family="ALU"
25798..26082
repeat_region /rpt_family="ALU"
repeat_region complement(27323..27403)
repeat_region /rpt_family="ALU"
27765..28012
repeat_region /rpt_family="ALU"
repeat_region complement(31012..31088)
repeat_region /rpt_family="ALU"
repeat_region complement(31600..31987)
repeat_region /rpt_family="L1"
repeat_region complement(32098..32504)
repeat_region /rpt_family="L1"
repeat_region complement(34329..34608)
repeat_region /rpt_family="ALU"
repeat_region 37299..37709
repeat_region /rpt_family="L1"
BASE COUNT 10675 a 7940 c 8152 g 11973 t
ORIGIN
Query Match 86.7%; Score 44.2; DB 9; Length 38740;
Best Local Similarity 90.2%; Pred. No. 4.8e-05;
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51
|||||
Db 15835 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 15785
|||||
RESULT 8
AC022117 69185 bp DNA linear HTG 18-JUL-2000
LOCUS
DEFINITION Homo sapiens chromosome 5 clone CTC-509C13, WORKING DRAFT SEQUENCE,
9 ordered pieces.
ACCESSION AC022117
VERSION AC022117.3 GI:7711695
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 69185)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 5
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 69185)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-2000) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On May 6, 2000 this sequence version replaced gi:7239444.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 380697, H426
Center clone name: CIT-HSPC_509C13
-----
Summary Statistics
Consensus quality: 56011 bases at least Q40
Consensus quality: 64421 bases at least Q30
Consensus quality: 66686 bases at least Q20
Estimated insert size: 70000; pulse field gel estimation
Estimated insert size: 68835; sum-of-contigs estimation
Quality coverage: 3.97 in Q20 bases; pulse field gel estimation
Quality coverage: 4.04 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 2503: contig of 2503 bp in length
* 2504 2603: gap of unknown length
* 2604 16029: contig of 13426 bp in length
* 16030 16129: gap of unknown length
* 16130 23795: contig of 7666 bp in length
* 23796 23895: gap of unknown length
* 23896 32869: contig of 8974 bp in length
* 32870 32969: gap of unknown length
* 32970 43827: contig of 10858 bp in length
* 43828 43927: gap of unknown length
* 43928 47286 47385: gap of unknown length
* 47386 51095: contig of 3710 bp in length
* 51096 51195: gap of unknown length
* 51196 67913: contig of 16718 bp in length
* 67914 68013: gap of unknown length
* 68014 69185: contig of 1172 bp in length.
FEATURES
source
1..69185
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-509C13"
/clone_lib="CalTech human BAC library C"
BASE COUNT 21070 a 12867 c 11934 g 22513 t 801 others
ORIGIN
Query Match 86.7%; Score 44.2; DB 2; Length 69185;
Best Local Similarity 90.2%; Pred. No. 4.6e-05;
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51
|||||
Db 8502 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 8552
|||||
RESULT 9
AC010222 86424 bp DNA linear HTG 18-JUL-2000
LOCUS
DEFINITION Homo sapiens chromosome 5 clone CTC-265B3, WORKING DRAFT SEQUENCE,
8 ordered pieces.
ACCESSION AC010222
VERSION AC010222.4 GI:9256163
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Homo sapiens.
```

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 86424)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Sequencing of Human Chromosome 5  
 JOURNAL Unpublished  
 2 (bases 1 to 86424)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Jul 18, 2000 this sequence version replaced gi:7710836.  
 -----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: http://www.jgi.doe.gov  
 -----  
 Project Information  
 Center Project Name: 286966  
 Center clone name: CIT-HSPC\_265B3  
 -----  
 Summary Statistics  
 Consensus quality: 81824 bases at least Q40  
 Consensus quality: 85101 bases at least Q30  
 Consensus quality: 85349 bases at least Q20  
 Estimated insert size: 85000; pulse field gel estimation  
 Estimated insert size: 86124; sum-of-contigs estimation  
 Quality coverage: 8.75 in Q20 bases; pulse field gel estimation  
 Quality coverage: 8.64 in Q20 bases; sum-of-contigs estimation.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 11208: contig of 11208 bp in length  
 \* 11308: gap of unknown length  
 \* 11309 16728: contig of 5420 bp in length  
 \* 16729 16828: gap of unknown length  
 \* 16829 32348: contig of 15520 bp in length  
 \* 32349 32448: gap of unknown length  
 \* 32449 59416: contig of 26968 bp in length  
 \* 59417 59516: gap of unknown length  
 \* 59517 70876: contig of 11360 bp in length  
 \* 70877 70976: gap of unknown length  
 \* 70977 75433: contig of 4457 bp in length  
 \* 75434 75533: gap of unknown length  
 \* 75534 79249: contig of 3716 bp in length  
 \* 79250 79350: gap of unknown length  
 \* 79350 86424: contig of 7075 bp in length.  
 \* Location/Qualifiers  
 1..86424  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="CTC-265B3"  
 /clone\_lib="Caltech human BAC library C"  
 BASE COUNT 26427 a 15999 c 15818 g 27477 t 703 others  
 ORIGIN  
 Query Match 86.7%; Score 44.2; DB 2; Length 86424;  
 Best Local Similarity 90.2%; Pred. No. 4.5e-05;  
 Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 TCAAGACACGCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51  
 ||||||||||||||||| ||| ||||| |||||||||||||||||  
 Db 50454 TCAAGACACGCTGGACAACTGGTAAACCCGCTCTCTACAAAAATACA 50504

RESULT 10  
 AL356748  
 LOCUS  
 DEFINITION Human DNA sequence from clone RP4-686J16 on chromosome 1, complete  
 sequence.  
 ACCESSION AL356748  
 VERSION AL356748  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 97559)  
 AUTHORS Laird, G.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-DEC-2000) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 COMMENT requests: clonerequest@sanger.ac.uk  
 On Dec 5, 2000 this sequence version replaced gi:11493283.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence has been finished according to sequence map criteria  
 as follows. An attempt is made to resolve all sequencing problems,  
 such as compressions and repeats, but not necessarily within known  
 annotated repeat sequence elements. Where the sequence is  
 ambiguous, there is an annotation using the 'unsure' feature key.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em.: EMBL; Sw.: SWISSPROT; Tr.: TREMBL; Wp.: WORMPEP; Information  
 on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/projects/C-elegans/wormpep/ This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping  
 Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chrl  
 RP4-686J16 is from the library RPCI-4 constructed by the group of  
 Plietser de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm  
 VECTOR: pCYPAC2  
 IMPORTANT: This sequence is not the entire insert of clone  
 RP4-686J16. It may be shorter because we sequence overlapping  
 sections only once, except for a 100 base overlap.  
 The true left end of clone RP4-686J16 is at 1 in this sequence. The  
 true left end of clone RP4-655N15 is at 97460 in this sequence. The  
 true right end of clone RP4-781D12 is at 1573 in this sequence.  
 Location/Qualifiers  
 1..97559  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /clone="RP4-686J16"  
 /clone\_lib="RPCI-4"  
 378..514  
 /note="MIR repeat: matches 81..226 of consensus"  
 927..1084  
 /note="MIR repeat: matches 100..258 of consensus"  
 1112..1353  
 /note="AluX repeat: matches 1..256 of consensus"  
 1512..1812  
 /note="AluX repeat: matches 1..300 of consensus"  
 complement(1745..2002)  
 /note="match: GSS: Em:AQ321132"  
 1820..1927  
 /note="AluJ repeat: matches 191..298 of consensus"  
 2008..2639  
 /note="match: GSS: Em:AQ474986"  
 2014..2416  
 /note="match: GSS: Em:AQ706294"  
 2051..2205

repeat\_region  
 repeat\_region  
 repeat\_region  
 repeat\_region  
 misc\_feature  
 repeat\_region  
 misc\_feature  
 misc\_feature  
 repeat\_region

```
repeat_region /note="MIR repeat: matches 78. .252 of consensus"
2328. .2414
repeat_region /note="MTLH repeat: matches 51. .130 of consensus"
2477. .2555
repeat_region /note="MTLH repeat: matches 202. .288 of consensus"
2676. .2813
misc_feature /note="MTLH repeat: matches 401. .534 of consensus"
complement(2908. .3284)
repeat_region /note="match: STS: Em:G18185"
3270. .3365
repeat_region /note="L2 repeat: matches 2075. .2178 of consensus"
4118. .4212
misc_feature /note="LIP3 repeat: matches 5725. .5822 of consensus"
7092. .7341
misc_feature /note="match: STS: Em:G15587"
complement(7568. .8088)
misc_feature /note="match: GSS: Em:AQ0707880"
complement(7909. .8281)
repeat_region /note="match: GSS: Em:AQ035467"
8073. .8228
repeat_region /note="MIR repeat: matches 106. .245 of consensus"
8286. .8342
repeat_region /note="L2 repeat: matches 2691. .2749 of consensus"
8660. .8744
repeat_region /note="L2 repeat: matches 2217. .2298 of consensus"
8745. .9049
repeat_region /note="AluY repeat: matches 1. .305 of consensus"
9050. .9230
misc_feature /note="L2 repeat: matches 2023. .2217 of consensus"
complement(10541. .11084)
misc_feature /note="match: GSS: Em:AQ316052"
complement(10633. .10966)
repeat_region /note="match: GSS: Em:A2121092"
11304. .11419
repeat_region /note="LTR9B repeat: matches 445. .560 of consensus"
12397. .12857
misc_feature /note="MTLH repeat: matches 31. .512 of consensus"
complement(13076. .13897)
repeat_region /note="match: GSS: Em:AQ743494"
13492. .13591
repeat_region /note="L2 repeat: matches 2388. .2487 of consensus"
13685. .13725
repeat_region /note="L2 repeat: matches 2709. .2749 of consensus"
14198. .14273
repeat_region /note="L2 repeat: matches 2673. .2749 of consensus"
15025. .15148
repeat_region /note="L2 repeat: matches 2584. .2708 of consensus"
16477. .16612
misc_feature /note="MIR repeat: matches 85. .228 of consensus"
complement(16832. .17296)
repeat_region /note="match: GSS: Em:AQ700892"
17109. .17167
misc_feature /note="MER45B repeat: matches 437. .492 of consensus"
17344. .17892
repeat_region /note="match: GSS: Em:B50102"
17496. .17773
repeat_region /note="AluSq repeat: matches 37. .311 of consensus"
18470. .18669
repeat_region /note="t1MC/D repeat: matches 5423. .5612 of consensus"
18751. .18902
repeat_region /note="t1MC5 repeat: matches 7758. .7913 of consensus"
18928. .19233
repeat_region /note="AluX repeat: matches 5. .312 of consensus"
19235. .19296
repeat_region /note="31 copies 2 mer aa 74% conserved"
19553. .19654
repeat_region /note="AluJo/FRAM repeat: matches 203. .304 of consensus"
19692. .19772
repeat_region /note="AluJ/FLAM repeat: matches 5. .85 of consensus"
20881. .20991
repeat_region /note="MIR repeat: matches 16. .142 of consensus"
21460. .21617
repeat_region /note="MIR repeat: matches 20. .184 of consensus"
```

```
repeat_region 22616. .23224
/note="LIME1 repeat: matches 5528. .6159 of consensus"
23288. .23435
repeat_region /note="MIR repeat: matches 80. .233 of consensus"
25186. .25354
repeat_region /note="MIR repeat: matches 77. .262 of consensus"
25530. .25827
misc_feature /note="AluX repeat: matches 1. .298 of consensus"
complement(26592. .27148)
misc_feature /note="match: GSS: Em:AQ532789"
complement(26598. .27124)
misc_feature /note="match: GSS: Em:AQ619294"
complement(26720. .27017)
misc_feature /note="match: GSS: Em:AQ100816"
complement(26833. .27131)
repeat_region /note="match: GSS: Em:AQ837606"
26971. .27041
misc_feature /note="MTLJ repeat: matches 110. .180 of consensus"
27111. .27147
misc_feature /note="Single clone region. Assembly confirmed by restriction digest data."
27131. .27224
misc_feature /note="match: GSS: Em:AQ136115"
27133. .27439
misc_feature /note="match: GSS: Em:AQ556313"
27139. .27631
repeat_region /note="match: GSS: Em:AQ798882"
27202. .27252
misc_feature /note="3 copies 17 mer 86% conserved"
27218. .27578
misc_feature /note="match: GSS: Em:AQ438628"
27236. .27568
misc_feature /note="match: GSS: Em:AQ136115"
29394. .29510
repeat_region /note="L2 repeat: matches 2587. .2699 of consensus"
29562. .29638
repeat_region /note="L2 repeat: matches 2587. .2664 of consensus"
32900. .32937
misc_feature /note="L2 repeat: matches 2660. .2705 of consensus"
33091. .33356
repeat_region /note="match: GSS: Em:AQ316053"
33999. .34281
repeat_region /note="AluY repeat: matches 17. .311 of consensus"
34351. .34640
repeat_region /note="AluX repeat: matches 1. .287 of consensus"
35737. .36047
repeat_region /note="AluSq repeat: matches 1. .313 of consensus"
37307. .37627
repeat_region /note="AluJo repeat: matches 1. .293 of consensus"
38057. .38173
repeat_region /note="FLAM.C repeat: matches 1. .118 of consensus"
38174. .38190
repeat_region /note="Charlie4 repeat: matches 1855. .1871 of consensus"
38191. .38449
repeat_region /note="AluX repeat: matches 1. .302 of consensus"
38450. .38544
```

Query Match 86.7%; Score 44.2; DB 9; Length 97559;  
Best Local Similarity 90.2%; Pred.No. 4.5e-05;  
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGCAACTGGAGAACCGGTCTCTACAAAAATACA 51  
|||||  
Db 65762 TCAAGACCGCTGGCAACATGTTAGACCCGCTCTACAAAAATACA 65812

RESULT 11  
AC091816/c  
LOCUS AC091816 97860 bp DNA linear PRI 07-MAR-2002  
DEFINITION Homo sapiens l2p BAC RP11-525I3 (Roswell Park Cancer Institute  
Human BAC Library) complete sequence.  
ACCESSION AC091816  
VERSION AC091816.8 GI:19033408

## KEYWORDS

SOURCE Homo sapiens.

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 97860)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C., Alsbrooks,S.L., Amaraturge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cannon,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Denny,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.R., Delgado,O., Duggan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,P., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homsif,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marandel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neale,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokoko,S., Ogun,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,D., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tanton,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalob,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.

Direct Submission

Unpublished

2 (bases 1 to 97860)

Worley,K.C.

Direct Submission

Submitted (09-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 97860)

Worley,K.C.

Direct Submission

Submitted (01-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 97860)

Worley,K.C.

Direct Submission

Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Mar 1, 2002 this sequence version replaced gi:15625994.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

## TITLE

JOURNAL

REFERENCE

AUTHORS

## JOURNAL

Submitted (09-JUN-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 97860)

Worley,K.C.

## REFERENCE

AUTHORS

## JOURNAL

Submitted (01-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 97860)

Worley,K.C.

## REFERENCE

AUTHORS

## JOURNAL

Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Mar 1, 2002 this sequence version replaced gi:15625994.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the

entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT.

FEATURES	source	Location/Qualifiers
		1..97860
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="12"
		/clone="RP11-52513"
		complement(1..208)
		/note="overlaps bases 1..208 of clone AC007537"
		/function="clone overlap"
		1..144
		/rpt_family="Tigger3(Golem)"
		complement(245..552)
		/rpt_family="AluJb"
		1146..1191
		/rpt_family="L2"
		complement(1442..1585)
		/rpt_family="MIR"
		complement(1586..1995)
		/rpt_family="MSTA"
		2041..2185
		/rpt_family="L2"
		complement(2584..2884)
		/rpt_family="AluSg"
		complement(3260..3410)
		/rpt_family="L1MEC"
		complement(3459..4286)
		/rpt_family="L1MEC"
		complement(4287..4578)
		/rpt_family="AluX"
		complement(4579..5154)
		/rpt_family="L1MEC"
		complement(5155..5457)
		/rpt_family="AluJo"
		complement(5458..5880)
		/rpt_family="L1MEC"
		complement(5902..5992)
		/rpt_family="MIR"
		6239..6325
		/rpt_family="L2"



```

repeat_region /note="MSTC element fragment"
22918..23211
repeat_region /note="THE1B element fragment"
23079..23258
repeat_region /note="MSTA element fragment"
23384..23602
repeat_region /note="L1 element fragment"
23772..23957
repeat_region /note="L1 element fragment"
24136..24249
repeat_region /note="L1 element fragment"
24312..25217
repeat_region /note="L1 element fragment"
26752..27026
repeat_region /note="MER42A element fragment"
26776..26875
repeat_region /note="L1 element fragment"
27831..28139
repeat_region /note="Alu repeat: matches 1..308 of consensus"
29106..29376
repeat_region /partial
29400..29687
repeat_region /note="Alu repeat: matches 308..1 of consensus"
29927..30033
repeat_region /note="MLT1A element fragment"
29943..30005
repeat_region /note="MSTC element fragment"
30431..30722
repeat_region /note="Alu repeat: matches 1..308 of consensus"
32718..32789
repeat_region /note="MSTA element fragment"
32733..32774
repeat_region /note="THE1B element fragment"
32846..32896
repeat_region /note="THE1B element fragment"
32923..33068
repeat_region /note="MSTA element fragment"
32961..33043
repeat_region /note="THE1B element fragment"
33360..41487
repeat_region /note="L1 element fragment"
41609..41727
repeat_region /note="MLT1D element fragment"
42134..42220
repeat_region /note="MLT1E element fragment"
42604..44783
repeat_region /note="L1 element fragment"
44819..45081
repeat_region /partial
repeat_region /note="Alu repeat: matches 21..308 of consensus"
45595..45894
repeat_region /partial
repeat_region /note="Alu repeat: matches 308..1 of consensus"
46842..47139
repeat_region /partial
repeat_region /note="Alu repeat: matches 308..1 of consensus"
48908..49199
repeat_region /note="Alu repeat: matches 1..308 of consensus"
50812..50883
repeat_region /note="MLT1B element fragment"
51397..51711
repeat_region /note="MLT1A element fragment"
51835..52808
repeat_region /note="L1 element fragment"
52869..52993
repeat_region /partial
repeat_region /note="Alu repeat: matches 8..145 of consensus"
52994..53107
repeat_region /partial
repeat_region /note="Alu repeat: matches 198..308 of consensus"
53138..53218

```

```

repeat_region /note="L1 element fragment"
53226..53363
repeat_region /partial
repeat_region /note="Alu repeat: matches 1..144 of consensus"
55035..55097
repeat_region /note="MLT1C element fragment"
55176..55292
repeat_region /note="MLT1C element fragment"
55194..55292
repeat_region /note="MLT1B element fragment"
56882..56944
repeat_region /note="MER1B element fragment"
56975..57263
repeat_region /partial
repeat_region /note="Alu repeat: matches 308..1 of consensus"
57291..57427
repeat_region /note="MER1B element fragment"
57354..57428
repeat_region /note="MER1A element fragment"
57504..57783
repeat_region /note="MER1B element fragment"
59781..60056
repeat_region /partial
repeat_region /note="Alu repeat: matches 306..1 of consensus"
60682..60973
repeat_region /note="Alu repeat: matches 1..308 of consensus"
63108..63158
repeat_region /note="MLT2C2 element fragment"
63339..63377
repeat_region /note="MLT2C2 element fragment"
64695..64823
repeat_region /note="L1 element fragment"
65160..65612
repeat_region /note="L1 element fragment"
65665..67096
repeat_region /note="L1 element fragment"
67152..67235
repeat_region /note="L1 element fragment"
67286..68295
repeat_region /note="L1 element fragment"
68390..68571
repeat_region /note="L1 element fragment"

```

Query Match 86.7%; Score 44.2; DB 9; Length 101076;  
 Best Local Similarity 90.2%; Pred. No. 4.5e-05;  
 Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAGACGAGCTGGACAACCTTGGGAAGACCGGGTCTCTACAAAAATACA 51  
 |||||  
 Db 29298 TCAAGACGAGCTGGACAACATGCAAAACCTCTCTACAAAAATACA 29248

RESULT 13  
 AC015669/c AC015669 106314 bp DNA linear PRI 01-MAY-2002  
 LOCUS  
 DEFINITION Homo sapiens chromosome, clone RP11-8L20, complete sequence.  
 AC015669  
 VERSION AC015669.12 GI:20377024  
 KEYWORDS HTG.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 106314)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,  
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,  
 Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,



Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lehoczyk,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Pcterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 106314)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,  
Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,  
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Olivier,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (19-APR-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 106314)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,  
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,  
Landers,T., Lehoczyk,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,  
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Olivier,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Testfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (01-MAY-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 1, 2002 this sequence version replaced gi:20198600.  
All repeats were identified using RepeatMasker:  
Smith, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: LI284  
Center clone name: 8\_L\_20  
-----

Only the first 106,3 kilobases of this clone are being submitted.  
The remainder overlaps accession number AC016866 [WICGR project  
L3452].

FEATURES	Location/Qualifiers
Source	1..106314
	/organism:"Homo sapiens"
	/db_xref="taxon:9606"
	/clone="RP11-8L20"
	/clone_lib="RPC1-11 Human Male BAC"
	complement(1..60)
repeat_region	/rpt_family:"MER41A"
repeat_region	61..364
repeat_region	/rpt_family:"AluY"
repeat_region	complement(365..803)
repeat_region	/rpt_family:"MER41A"
repeat_region	complement(2476..2522)
repeat_region	/rpt_family:"L2"
repeat_region	2660..2864
repeat_region	/rpt_family:"MIR"
repeat_region	3113..3200
repeat_region	/rpt_family:"(TA)n"
repeat_region	complement(3428..3738)
repeat_region	/rpt_family:"AluY"
repeat_region	3750..3836
repeat_region	/rpt_family:"LMB3"
repeat_region	3900..3950
repeat_region	/rpt_family:"(CA)n"
repeat_region	3951..4123
repeat_region	/rpt_family:"LMB3"
repeat_region	4124..4417
repeat_region	/rpt_family:"AluY"
repeat_region	4418..4598
repeat_region	/rpt_family:"LMB3"
repeat_region	complement(4599..4701)
repeat_region	/rpt_family:"AluSg/x"
repeat_region	4810..4842
repeat_region	/rpt_family:"Alu"
repeat_region	4876..4967
repeat_region	/rpt_family:"MIR"
repeat_region	4996..5228
repeat_region	/rpt_family:"MIR"
repeat_region	5245..5380
repeat_region	/rpt_family:"AluSg"
repeat_region	5381..5683
repeat_region	/rpt_family:"AluY"
repeat_region	5684..5842
repeat_region	/rpt_family:"AluSg"
repeat_region	5845..5867
repeat_region	/rpt_family:"AT-rich"
repeat_region	complement(6025..6163)
repeat_region	/rpt_family:"L3"
repeat_region	complement(6164..6451)
repeat_region	/rpt_family:"AluJb"
repeat_region	complement(6452..6688)
repeat_region	/rpt_family:"L3"
repeat_region	6727..6886
repeat_region	/rpt_family:"GA-rich"
repeat_region	7241..7374
repeat_region	/rpt_family:"MLT1"
repeat_region	7441..7550
repeat_region	/rpt_family:"MLT1"
repeat_region	complement(7721..8025)
repeat_region	/rpt_family:"AluJb"
repeat_region	8118..8226
repeat_region	/rpt_family:"L1MD2"
repeat_region	8571..8882
repeat_region	/rpt_family:"AluY"
repeat_region	complement(9103..9412)

```

/rpt_family="AluSg1"
complement(9436..9743)
/rpt_family="AluSx"
complement(9790..10013)
/rpt_family="L1ME3"
complement(10014..10309)
/rpt_family="AluSg1"
complement(10629..11046)
/rpt_family="L1MC3"
complement(11079..11151)
/rpt_family="MIR1F1"
complement(11180..11286)
/rpt_family="L1MD3"
complement(11287..11597)
/rpt_family="AluSg"
complement(11598..11818)
/rpt_family="L1MD3"
12224..12271
/rpt_family="(CA)n"
complement(13320..13398)
/rpt_family="MIR"
14195..14216
/rpt_family="GC_rich"
14267..14328
/rpt_family="GC_rich"
14363..14390
/rpt_family="GC_rich"
14480..14517
/rpt_family="GC_rich"
14629..14671
/rpt_family="GC_rich"
15073..15396
/rpt_family="AluSx"
15610..15855
/rpt_family="L1ME2"

Query Match      86.7%; Score 44.2; DB 9; Length 106314;
Best Local Similarity 90.2%; Pred. No. 4.5e-05;
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGCAACTGGAGACCGTCTCTACAAAAATACA 51
|||||
Db 102175 TCAAGACCGCTGGCAACTGGAGACCGTCTCTACAAAAATACA 102125
|||||

RESULT 14
AL133461
LOCUS      AL133461      124328 bp      DNA      linear      PRI 16-NOV-2001
DEFINITION Human DNA sequence from clone RP11-359E7 on chromosome 10, complete
sequence.
ACCESSION  AL133461
VERSION     AL133461.10 GI:16973782
KEYWORDS   HTG.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 124328)
AUTHORS   Wall,M.
TITLE     Direct Submission
JOURNAL   Submitted (16-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
hamquery@sanger.ac.uk
On Nov 17, 2001 this sequence version replaced gi:14970653.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone.
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=

/rpt_family="AluSg1"
complement(9436..9743)
/rpt_family="AluSx"
complement(9790..10013)
/rpt_family="L1ME3"
complement(10014..10309)
/rpt_family="AluSg1"
complement(10629..11046)
/rpt_family="L1MC3"
complement(11079..11151)
/rpt_family="MIR1F1"
complement(11180..11286)
/rpt_family="L1MD3"
complement(11287..11597)
/rpt_family="AluSg"
complement(11598..11818)
/rpt_family="L1MD3"
12224..12271
/rpt_family="(CA)n"
complement(13320..13398)
/rpt_family="MIR"
14195..14216
/rpt_family="GC_rich"
14267..14328
/rpt_family="GC_rich"
14363..14390
/rpt_family="GC_rich"
14480..14517
/rpt_family="GC_rich"
14629..14671
/rpt_family="GC_rich"
15073..15396
/rpt_family="AluSx"
15610..15855
/rpt_family="L1ME2"

Query Match      86.7%; Score 44.2; DB 9; Length 106314;
Best Local Similarity 90.2%; Pred. No. 4.5e-05;
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGCAACTGGAGACCGTCTCTACAAAAATACA 51
|||||
Db 102175 TCAAGACCGCTGGCAACTGGAGACCGTCTCTACAAAAATACA 102125
|||||

RESULT 15
AC026773
LOCUS      AC026773      130965 bp      DNA      linear      PRI 16-NOV-2001
DEFINITION Homo sapiens chromosome 5 clone CTC-230L18, complete sequence.
ACCESSION  AC026773
VERSION     AC026773.6 GI:16924092
KEYWORDS   HTG.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 130965)
AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE     Direct Submission
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 130965)
AUTHORS   DOE Joint Genome Institute.
TITLE     Direct Submission
JOURNAL   Submitted (24-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE  3 (bases 1 to 130965)
AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE     Direct Submission
JOURNAL   Submitted (01-MAR-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE  4 (bases 1 to 130965)
AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE     Direct Submission
JOURNAL   Submitted (15-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE  5 (bases 1 to 130965)

30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormep/
This sequence was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-359E7 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-359E7 It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true right end of clone RP11-359E7 is at 124328 in this
sequence. The true right end of clone RP11-88110 is at 2000 in this
sequence.

FEATURES             Location/Qualifiers
     Source           1..124328
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-359E7"
                     /clone_lib="RPCI-11.2"
BASE COUNT  36982 a 24571 c 26217 g 36558 t
ORIGIN
Query Match      86.7%; Score 44.2; DB 9; Length 124328;
Best Local Similarity 90.2%; Pred. No. 4.4e-05;
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAGACCGCTGGACACTGGAGACCGTCTCTACAAAAATACA 51
|||||
Db 109810 TCAGACCGCTGGACACTGGAGACCGTCTCTACAAAAATACA 109860
|||||

RESULT 15
AC026773
LOCUS      AC026773      130965 bp      DNA      linear      PRI 16-NOV-2001
DEFINITION Homo sapiens chromosome 5 clone CTC-230L18, complete sequence.
ACCESSION  AC026773
VERSION     AC026773.6 GI:16924092
KEYWORDS   HTG.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 68364)
AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE     Direct Submission
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 130965)
AUTHORS   DOE Joint Genome Institute.
TITLE     Direct Submission
JOURNAL   Submitted (24-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE  3 (bases 1 to 130965)
AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE     Direct Submission
JOURNAL   Submitted (01-MAR-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE  4 (bases 1 to 130965)
AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE     Direct Submission
JOURNAL   Submitted (15-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE  5 (bases 1 to 130965)
```

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (14-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
REFERENCE 6 (bases 1 to 130965)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Nov 14, 2001 this sequence version replaced gi:15187257.  
FEATURES  
source  
1..130965  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTC-230L18"  
complement(1..68364)  
misc\_feature  
/note="Draft Sequence Produced by DOE Joint Genome  
Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 99.8% of Sequence;  
Estimated Total Number of Errors is 0.3."  
BASE COUNT 41144 a 24526 c 23780 g 41515 t  
ORIGIN  
Query Match 86.7%; Score 44.2; DB 9; Length 130965;  
Best Local Similarity 90.2%; Pred. No. 4.4e-05;  
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 TCAAGACCAGCCTGGACAACCTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 92661 TCAAGACCAGCCTGGACAACATGGTAAACCCGGTCTCTACAAAAATACA 92711  
|||||  
Search completed: February 11, 2003, 05:28:18  
Job time : 293.499 secs





```

/*tag= f
/Note= "sequencing primer F3"
1124..1143
/*tag= g
/Note= "sequencing primer R4"
1605..1623
/*tag= h
/Note= "sequencing primer R5"
1650..1669
/*tag= i
/Note= "amplification primer uprl669"
1532..1619
/*tag= j
1620..1669
/*tag= k
/partial
/Note= "5' region of CYP2D6 coding region"

```

WO200155432-A2.

02-AUG-2001.

30-JAN-2001; 2001WO-EP00954.

31-JAN-2000; 2000EP-0101889.

(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

Raimundo S, Zanger U;

WPI; 2001-457734/49.

A polynucleotide capable of hybridizing to CYP2D6 promoter useful for the optimization of drug therapies using substrates of cytochrome p-450

Claim 1; Fig 1; 41pp; English.

The present sequence is that of the promoter region of the human cytochrome P450 CYP2D6 gene. The promoter region was amplified by PCR from leucocyte DNA of over 50 individuals, and sequenced. 8 Previously unknown single nucleotide polymorphisms (SNP) were identified. These were at: Base 36 (base -1584 according to the Human Cytochrome P450 Allele Nomenclature), where the SNP was C to G, occurring at an estimated frequency of approximately 20% in the whole population, and resulting in increased enzyme activity; position 194 (-1426), C to T, approximately 20% frequency; neutral function; position 385 (-1235), A to G, approximately 50% frequency; neutral function; position 620 (-1000), G to A, approximately 20% frequency; neutral function; position 880 (-740), C to T, approximately 30% frequency; unknown function; position 940 (-680), G to A, approximately 30% frequency; unknown function; 1255 (-365), G to A, rare, unknown function; and 1298 (-322), T to C, rare, unknown function. The C to G mutation at -1584 bp is strongly associated with lower metabolic ratios, and a molecular variant Polynucleotide having G at this position is claimed (see AAH26179). The invention provides a method of diagnosing a disorder related to reduced or enhanced capacity for clearance of CYP2D6 substrates (antiarrhythmic, beta-adrenergic receptor-antagonist, tricyclic antidepressant, selective serotonin reuptake inhibitor, tricyclic opiate, cytotatic or amphetamine), or opiate, cytotatic or amphetamine), or susceptibility to such a disorder, by determining the presence of the mutation. The strong association of the common C to G mutation at -1584 bp with increased enzyme activity significantly improves the correlation between genotype and phenotype in the CYP2D6 polymorphism. Testing for the mutation will allow the identification of intermediate metabolizers and therefore allow quantitative predictions to be made on in vivo drug metabolism capacity, thus providing a very potent tool for improving the therapy of diseases with drugs that are targets of the CYP2D6 gene product.

Sequence 1669 BP; 413 A; 376 C; 534 G; 338 T; 8 other;

Query Match 99.2%; Score 50.6; DB 22; Length 1669;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACTTGGAGAAGACCGGTCCTCTACAAAAATACA 51  
 |||||  
 Db 5 TCAAGACCAGCCTGGACAACTTGGAGAAGACCGGTCCTCTACAAAAATACA 55

RESULT 2

AAH26179

ID AAH26179 standard; DNA; 1669 BP.

XX AC AAH26179;

XX DT 17-SEP-2001 (first entry)

XX DE Human cytochrome P450 CYP2D6 gene promoter (G mutation at -1584 bp).

XX KW Cytochrome P450; CYP2D6; promoter; drug metabolism; human;  
 diagnosis; therapy; single nucleotide polymorphism; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT variation replace(36,G)  
 FT /\*tag= a

FT /\*frequency= "20%"

FT 5'UTR 1532..1619

FT /\*tag= b

FT CDS 1620..1669

FT /\*tag= c

FT /\*partial

FT /\*note= "5' region of CYP2D6 coding region"

XX PN WO200155432-A2.

XX PD 02-AUG-2001.

XX PD 30-JAN-2001; 2001WO-EP00954.

XX PR 31-JAN-2000; 2000EP-0101889.

XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX PI Raimundo S, Zanger U;

XX WPI; 2001-457734/49.

A polynucleotide capable of hybridizing to CYP2D6 promoter useful for the optimization of drug therapies using substrates of cytochrome P-450

Claim 1(a); Page -; 41pp; English.

The present sequence is that of the promoter region of the human cytochrome P450 CYP2D6 gene, which includes G at position 36 (base -1584 according to the Human Cytochrome P450 Allele Nomenclature). The presence of C at position -1584 bp is a marker for low enzyme activity, whereas there is strong association of G at position -1584 bp with increased enzyme activity. The C to G single nucleotide polymorphism occurs in approximately 20% of the population. The invention provides a method of diagnosing a disorder related to reduced or enhanced capacity for clearance of CYP2D6 substrates (antiarrhythmic, beta-adrenergic receptor antagonist, tricyclic antidepressant, selective serotonin reuptake inhibitor, neuroleptic, opiate, cytotatic or amphetamine), or susceptibility to such a disorder, by determining the presence of a mutation in the CYP2D6 promoter. The novel variant forms of the CYP2D6 gene provided by the invention provide the potential for the development of a pharmacodynamic profile of drugs for a given patient. The finding and characterization of variations in the CYP2D6 gene, and diagnostic tests for the discrimination of

CC different alleles in human individuals, provide a very potent tool  
CC for improving the therapy of diseases with drugs that are targets  
CC of the CYP2D6 gene production, and whose metabolism is therefore  
CC dependent on CYP2D6 activity.  
CC Note: The present sequence is not shown in the specification but is  
CC derived from the CYP2D6 promoter sequence given in the Sequence  
CC Listing (see AH26169).

XX SQ Sequence 1669 BP; 413 A; 376 C; 535 G; 338 T; 7 other;  
Query Match 99.2%; Score 50.6; DB 22; Length 1669;  
Best Local Similarity 98.0%; Pred. No. 1.2e-09;  
Matches 50; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAAATACA 51  
|||||  
DB 5 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAAATACA 55

RESULT 3  
AAD34214  
ID AAD34214 standard; DNA; 1680 BP.  
XX AC AAD34214;  
XX DT 16-JUL-2002 (first entry)  
XX DE Human CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX KW Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic; 22;  
KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; polymorphism; ds.  
XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
FT misc\_feature 36  
FT /\*tag= a  
FT /\*note= "Polymorphic site"  
FT misc\_feature 194  
FT /\*tag= b  
FT /\*note= "Polymorphic site"  
FT misc\_feature 385  
FT /\*tag= c  
FT /\*note= "Polymorphic site"  
FT misc\_feature 620  
FT /\*tag= d  
FT /\*note= "Polymorphic site"  
FT misc\_feature 880  
FT /\*tag= e  
FT /\*note= "Polymorphic site"  
FT misc\_feature 942  
FT /\*tag= f  
FT /\*note= "Polymorphic site"  
FT misc\_feature 1255  
FT /\*tag= g  
FT /\*note= "Polymorphic site"

XX WO200218638-A2.  
XX PD 07-MAR-2002.  
XX PF 27-AUG-2001; 2001WO-IB01544.  
XX PR 30-AUG-2000; 2000GB-0021286.  
XX PA (GEMI-) GEMINI GENOMICS PLC.  
XX PI Risinger C, Andersson MK, Lewander T, Ollasson E;  
XX WPI; 2002-329785/36.  
XX DR New sequence determination oligonucleotides, useful for detecting  
XX PT

PT polymorphic sites in a 5' flanking region of a CYP2D6 gene, as  
PT hybridization probes, as components of diagnostic assays, or in  
PT ligase-based sequence determination -  
XX Claim 1; Fig 2; 63pp; English.  
XX CC The invention relates to sequence determination oligonucleotides for  
CC detecting polymorphic sites in a 5' flanking region of cytochrome P450  
CC 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many  
CC different xenobiotics. Human CYP2D6 gene is located on chromosome 22.  
CC The oligonucleotides may be used as in situ hybridisation probes, in  
CC ligase-based sequenced determination, as components of diagnostic assays,  
CC as probes in sequence determination methods based on mismatches, as  
CC hybridisation-based diagnostic assays, and as components of diagnostic  
CC microarray. CYP2D6 is useful to predict variations in an individual's  
CC ability to metabolise certain drugs. The present sequence is human  
CC CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX SQ Sequence 1680 BP; 413 A; 379 C; 539 G; 342 T; 7 other;

Query Match 99.2%; Score 50.6; DB 24; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 1.2e-09;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAAATACA 51  
|||||

DB 5 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAAATACA 55

RESULT 4  
AAD34213  
ID AAD34213 standard; DNA; 9432 BP.  
XX AC AAD34213;  
XX DT 16-JUL-2002 (first entry)  
XX DE Human cytochrome P450 2D6 (CYP2D6) gene.  
XX KW Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; ds.  
XX OS Homo sapiens.  
XX PN WO200218638-A2.  
XX PD 07-MAR-2002.  
XX PF 27-AUG-2001; 2001WO-IB01544.  
XX PR 30-AUG-2000; 2000GB-0021286.  
XX PA (GEMI-) GEMINI GENOMICS PLC.  
XX PI Risinger C, Andersson MK, Lewander T, Ollasson E;  
XX WPI; 2002-329785/36.

XX New sequence determination oligonucleotides, useful for detecting  
PT polymorphic sites in a 5' flanking region of a CYP2D6 gene, as  
PT hybridization probes, as components of diagnostic assays, or in  
PT ligase-based sequence determination -  
XX Example 3; Fig 1; 63pp; English.

XX CC The invention relates to sequence determination oligonucleotides for  
CC detecting polymorphic sites in a 5' flanking region of cytochrome P450  
CC 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many  
CC different xenobiotics. Human CYP2D6 gene is located on chromosome 22.  
CC The oligonucleotides may be used as in situ hybridisation probes, in  
CC ligase-based sequenced determination, as components of diagnostic assays,  
CC as probes in sequence determination methods based on mismatches, as

CC hybridisation-based diagnostic assays, and as components of diagnostic  
CC microarray. CYP2D6 is useful to predict variations in an individual's  
CC ability to metabolise certain drugs. The present sequence is human  
CC CYP2D6 gene.

XX  
SQ Sequence 9432 BP; 1964 A; 2647 C; 2976 G; 1845 T; 0 other;  
Query Match 99.2%; Score 50.6; DB 24; Length 9432;  
Best Local Similarity 98.0%; Pred. No. 1.9e-09;  
Matches 50; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCAAGACGAGCTGGACAACCTTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 5 TCAAGACGAGCTGGACAACCTTGGAGAACCCGGTCTCTACAAAAATACA 55

RESULT 5  
AAK96153/C  
ID AAK96153 standard; DNA; 401 BP.  
XX  
AC AAK96153;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE Human neuregulin gene single nucleotide polymorphism SNP8NRG1022025.  
XX  
KW Human; neuregulin-1 associated gene 1; NRGLAG1; Schizophrenia gene;  
KW gene therapy; single nucleotide polymorphism; SNP; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200164876-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-US06376.  
XX  
PR 28-FEB-2000; 2000US-0515715.  
XX  
PA (DECO-) DECODE GENETICS EHF.  
PI Stefansson H, Steinthorsdottir V, Gulcher JR;  
XX WPI; 2001-550179/61.  
XX  
PT Neuregulin-1 associated gene 1 nucleic acids and fragments, useful for  
PT preventing diagnosing and treating schizophrenia -  
XX  
PS Disclosure; Page 661; 750pp; English.  
XX  
CC This sequence represents a single nucleotide polymorphism (SNP) of the  
CC human neuregulin-1 associated gene 1 (NRGLAG1) of the invention. The  
CC NRGLAG1 gene is also referred to as the human Schizophrenia gene. The  
CC invention also relates to fragments or variants of the gene and the  
CC NRGLAG1 polypeptides they encode. The NRGLAG1 nucleic acids and  
CC polypeptides may be used in the prevention, diagnosis and treatment of  
CC diseases associated with inappropriate NRGLAG1 expression. For example,  
CC they may be used to treat disorders associated with decreased expression  
CC by rectifying mutations or deletions in a patient's genome that affect  
CC the activity of NRGLAG1 by expressing inactive proteins or to supplement  
CC the patients own production of NRGLAG1. Additionally, the gene may be  
CC used to produce NRGLAG1 polypeptides, by inserting the nucleic acids into  
CC a host cell and culturing the cell to express the protein. The gene may  
CC also be used as DNA probes and primers in diagnostic assays to detect and  
CC quantitate the presence of similar nucleic acids in samples, and  
CC therefore which patients may be in need of restorative therapy. The  
CC NRGLAG1 polypeptides may also be used as antigens in the production of  
CC antibodies against NRGLAG1 and in assays to identify modulators of  
CC NRGLAG1 expression and activity. Anti-NRGLAG1 antibodies and antagonists  
CC may also be used to down regulate expression and activity. Anti-NRGLAG1  
CC antibodies may also be used as diagnostic agents for detecting the  
CC presence of NRGLAG1 polypeptides in samples. NRGLAG1 is associated with  
CC schizophrenia which may be prevented, diagnosed and/or treated by the

CC above methods.  
XX  
SQ Sequence 401 BP; 107 A; 88 C; 93 G; 112 T; 1 other;  
Query Match 83.5%; Score 42.6; DB 22; Length 401;  
Best Local Similarity 88.2%; Pred. No. 9.4e-07;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACGAGCTGGACAACCTTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 84 TCAAGACGAGCTGGGCAACATGGAACCCCTGTCTCTACCAAAATACA 34

RESULT 6  
AAK97646/C  
ID AAK97646 standard; DNA; 401 BP.  
XX  
AC AAK97646;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE Human neuregulin gene single nucleotide polymorphism SNP8NRG1022025.  
XX  
KW Human; neuregulin 1 gene; schizophrenia; gene therapy; SNP;  
KW single nucleotide polymorphism; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200164877-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-US06377.  
XX  
PR 28-FEB-2000; 2000US-0515716.  
XX  
PA (DECO-) DECODE GENETICS EHF.  
PI Stefansson H, Steinthorsdottir V, Gulcher JR;  
XX WPI; 2001-514841/56.  
XX  
PT Neuregulin 1 nucleic acids and proteins useful for diagnosing  
PT preventing and treating schizophrenia -  
XX  
PS Disclosure; Page 246; 756pp; English.  
XX  
CC This sequence represents a single nucleotide polymorphism (SNP)  
CC from the human neuregulin 1 gene of the invention.  
CC The invention also relates to fragments or variants of the neuregulin 1  
CC gene. The gene and its proteins may be used in the prevention, diagnosis  
CC and treatment of diseases associated with inappropriate neuregulin 1  
CC expression, such as schizophrenia. For example they may be used to treat  
CC disorders associated with decreased neuregulin 1 expression by rectifying  
CC mutations or deletions in a patient's genome that affect the activity of  
CC neuregulin 1 by expressing inactive proteins or to supplement the  
CC patients own production of polypeptides. Additionally, the gene may be  
CC used to produce the neuregulin 1 protein, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the protein. The gene  
CC and its complementary sequences may also be used as DNA probes in  
CC diagnostic assays to detect and quantitate the presence of similar  
CC nucleic acids in samples, and therefore which patients may be in need of  
CC restorative therapy. The protein may also be used as antigens in the  
CC production of antibodies against neuregulin 1 and in assays to identify  
CC modulators of neuregulin 1 expression and activity. The antibodies and  
CC antagonists may also be used to down regulate expression and activity.  
CC The antibodies may also be used as diagnostic agents for detecting the  
CC presence of neuregulin 1 in samples.  
XX  
SQ Sequence 401 BP; 107 A; 88 C; 93 G; 112 T; 1 other;  
Query Match 83.5%; Score 42.6; DB 22; Length 401;  
Best Local Similarity 88.2%; Pred. No. 9.4e-07;



Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 84 TCAAGACCAGCCTGGGCAACATGGAAAAACCCTGTCTCTACCAAAAAATACA 34  
|||||

## RESULT 7

AAS65999/c  
ID AAS65999 standard; cDNA; 447 BP.

XX AAS65999;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #1803.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

XX P-PSDB; ABG01812.

XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX Claim 1: SEQ ID No 1803; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and

XX polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX and gene mapping, and in recombinant production of (II). The

XX polynucleotides are also used in diagnostics as expressed sequence tags

XX for identifying expressed genes. (I) is useful in gene therapy techniques

XX to restore normal activity of (II) or to treat disease states involving

XX (II). (II) is useful for generating antibodies against it, detecting or

XX quantitating a polypeptide in tissue, as molecular weight markers and as

XX a food supplement. (II) and its binding partners are useful in medical

XX imaging of sites expressing (II). (I) and (II) are useful for treating

XX disorders involving aberrant protein expression or biological activity.

XX The polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. AAS64197-AAS94564 represent novel human

XX diagnostic coding sequences of the invention.

XX Note: The sequence data for this patent did not appear in the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 447 BP: 84 A; 132 C; 101 G; 130 T; 0 other;

## Query Match

Best Local Similarity 83.5%; Score 42.6; DB 23; Length 447;

Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 177 TCAAGACCAGCCTGGGCAACATGGAGAACCCCTGTCTCTACAAAAATACA 127  
|||||

## RESULT 8

ABV52891

ID ABV52891 standard; cDNA; 556 BP.

XX AC ABV52891;

XX 17-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 52882.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

OS WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.

XX 17-FEB-2000; 2000US-183319P.

XX 16-MAR-2000; 2000US-189862P.

XX 25-MAY-2000; 2000US-207454P.

XX 09-JUN-2000; 2000US-211314P.

XX 18-JUL-2000; 2000US-219007P.

XX 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI: 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX Claim 1: Page 10252; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX specification or its complement. (I) is useful for:

XX (a) assessing whether a patient is afflicted with prostate cancer;

XX (b) monitoring the progression of prostate cancer in a patient;

XX (c) assessing the efficacy of a test compound to inhibit prostate

XX cancer in a patient;

XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer

XX in a patient;

XX (e) selecting a composition for inhibiting prostate cancer in a patient;

XX (f) assessing the prostate cell carcinogenic potential of a compound;

XX (g) determining whether prostate cancer has metastasized in a patient;

XX (h) assessing the aggressiveness or indolence of prostate cancer in a

XX patient;

XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 556 BP; 168 A; 119 C; 127 G; 141 T; 1 other;

## Query Match

Best Local Similarity 83.5%; Score 42.6; DB 23; Length 556;

Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||

Db 329 TCAAGACCAGCCTGGGCAACATGGCAAAACCCTGTCTCTACAAAAATACA 379  
|||||

## RESULT 9

```

ABN63076
ID  ABN63076 standard; cDNA; 565 BP.
XX
AC  ABN63076;
XX
DT  28-JUN-2002 (first entry)
XX
DE  Human cancer related polynucleotide SEQ ID NO 3043.
XX
KW  Human; cytostatic; gene expression; gene mapping; tissue profiling;
KW  gene therapy; cancer; tumour; gene; ss.
XX
OS  Homo sapiens.
XX
PN  WO200214500-A2.
XX
PD  21-FEB-2002.
XX
PF  16-AUG-2001; 2001WO-US25940.
XX
PR  16-AUG-2000; 2000US-226326P.
XX
PA  (CHIR ) CHIRON CORP.
PA  (HYSE-) HYSEQ INC.
XX
PI  Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
PI  Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX
DR  WPI; 2002-241905/29.
XX
PT  New nucleic acid for producing a polypeptide, detecting differentially
PT  expressed genes correlated with a cancerous state of a mammalian cell,
PT  and inhibiting tumor growth -
XX
PS  Claim 1; SEQ ID NO 3043; 883pp + Sequence Listing; English.
XX
CC  The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC  with cytostatic activity. The polynucleotide is used to produce a
CC  polypeptide, to detect differentially expressed genes correlated with a
CC  cancerous state of a mammalian cell and to inhibit tumour growth. The
CC  polynucleotide is used as a probe in mapping and tissue profiling. The
CC  encoded polypeptide and antibodies to the polypeptide can also be used
CC  for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC  gene therapy.
CC  Note: The sequence data for this patent did not form part of the printed
CC  specification, but was obtained in electronic format directly from WIPO
CC  at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ  Sequence 565 BP; 185 A; 96 C; 117 G; 167 T; 0 other;

    Query Match      83.5%; Score 42.6; DB 24; Length 565;
    Best Local Similarity 88.2%; Pred. No. 1e-06;
    Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY  1 TCAAGACCGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51
    |||||||
Db  437 TCAAGACCGCCTGGGCAACATGGCAAAACCCGTCTCTACAAAAATACA 487

RESULT 10
ABV50278
ID  ABV50278 standard; cDNA; 578 BP.
XX
AC  ABV50278;
XX
DT  17-SEP-2002 (first entry)
XX
DE  Human prostate expression marker cDNA 50269.
XX
KW  Human; prostate cancer; cytostatic; carcinogen; pharmacodynanamic marker;
KW  pharmacogenomic marker; gene; ss.
XX
OS  Homo sapiens.

```

```

XX
PN  WO200160860-A2.
XX
PD  23-AUG-2001.
XX
PF  20-FEB-2001; 2001WO-US05171.
XX
PR  17-FEB-2000; 2000US-183319P.
PR  16-MAR-2000; 2000US-189862P.
PR  25-MAY-2000; 2000US-207454P.
PR  09-JUN-2000; 2000US-211314P.
PR  18-JUL-2000; 2000US-219007P.
PR  13-DEC-2000; 2000US-255281P.
XX
PA  (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI  Schlegel R, Endege WO, Monahan JE;
XX
DR  WPI; 2001-662795/76.
XX
PT  Novel isolated nucleic acid molecule associated with cancerous state of
PT  prostate cells and correlating with presence of prostate cancer, useful
PT  for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS  Claim 1; Page 9788; 11750pp; English.
XX
CC  The invention relates to an isolated nucleic acid molecule (I) comprising
CC  a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC  specification or its complement. (I) is useful for:
CC  (a) assessing whether a patient is afflicted with prostate cancer;
CC  (b) monitoring the progression of prostate cancer in a patient;
CC  (c) assessing the efficacy of a test compound to inhibit prostate
CC  cancer in a patient;
CC  (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC  in a patient;
CC  (e) selecting a composition for inhibiting prostate cancer in a patient;
CC  (f) assessing the prostate cell carcinogenic potential of a compound;
CC  (g) determining whether prostate cancer has metastasized in a patient;
CC  (h) assessing the aggressiveness or indolence of prostate cancer in a
CC  patient;
CC  (i) is also useful as a pharmacodynanamic or pharmacogenomic marker.
XX
SQ  Sequence 578 BP; 167 A; 120 C; 120 G; 170 T; 1 other;

    Query Match      83.5%; Score 42.6; DB 23; Length 578;
    Best Local Similarity 88.2%; Pred. No. 1e-06;
    Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY  1 TCAAGACCGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51
    |||||||
Db  512 TCAAGACCGCCTGGACAACTGGCAAAACCCGTCTCTACTAAAAATACA 562

RESULT 11
AAS42150/c
ID  AAS42150 standard; DNA; 2268 BP.
XX
AC  AAS42150;
XX
DT  17-DEC-2001 (first entry)
XX
DE  Genomic sequence #466 encoding novel human enzyme polypeptide.
XX
KW  Human; oxidoreductase enzyme; transferase; hydrolase; lyase; isomerase;
KW  ligase; hyperproliferative disorder; immunodeficiency disorder;
KW  autoimmune disorder; neurological disorder; metabolic disorder;
KW  inflammatory disorder; cardiovascular disorder; reproductive disorder;
KW  blood-related disorder; infectious disorder; gene therapy; cytostatic;
KW  anti arthritic; nephrotropic; anticoagulant; ds.
XX
OS  Homo sapiens.
XX
PN  WO200155301-A2.

```

XX  
PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US011239.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180528.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 29-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0241827.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
(HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX





PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234423.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-451930/48.  
XX  
XX New cardiovascular system related polynucleotides and polypeptides,  
PT useful for diagnosing, treating and/or preventing disorders of the  
PT cardiovascular system -  
XX  
PS Claim 1; SEQ ID No 1902; 674pp; English.  
XX  
XX Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode  
CC the cardiovascular system antigen polypeptides of the invention.  
CC Cardiovascular system antigens and their associated polynucleotides are  
CC useful in the diagnosis, treatment and prevention of various types of  
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. A pathological condition can be determined by  
CC detecting the presence or absence of a mutation in a cardiovascular  
CC system antigen polynucleotide. The treatable disorders include autoimmune  
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such  
CC as neoplasms of the breast or liver, cardiovascular disorders such as  
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,  
CC nervous system disorders such as Alzheimer's disease, infections caused  
CC by bacteria, viruses and fungi, ocular disorders such as corneal

CC Infection, endocrine disorders such as premature labour and infertility,  
CC gastrointestinal disorders such as Crohn's disease, renal disorders such  
CC as glomerulonephritis and respiratory disorders such as asthma and  
CC pleurisy. The polypeptides can also be used to aid wound healing, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, to regenerate tissues and in chemotaxis.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

Query Match 83.5%; Score 42.6; DB 22; Length 4962;  
Best Local Similarity 88.2%; Pred. No. 1.9e-06;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TCAAGACGACCTGGACAACTTGAAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 2690 TCAAGACGACCTGGACAACTGGTGAACCCGTCTCTACAAAAATACA 2740  
|||||

RESULT 14  
AAS36403  
ID AAS36403 standard; DNA; 4962 BP.

XX AAS36403;

AC AAS36403;

XX DT 17-DEC-2001 (first entry)

XX DE Human cardiovascular system antigen genomic DNA SEQ ID No 1903.

XX KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-infertility.

XX OS Homo sapiens.

XX WO200155321-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01340.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 14-JUL-2000; 2000US-0217496.

PR 26-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246536.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 05-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-451930/48.  
XX  
XX New cardiovascular system related polynucleotides and polypeptides,  
XX useful for diagnosing, treating and/or preventing disorders of the  
XX cardiovascular system -  
XX  
XX Claim 1; SEQ ID NO 1903; 674pp; English.  
XX  
XX Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode  
XX the cardiovascular system antigen polypeptides of the invention.  
XX Cardiovascular system antigens and their associated polynucleotides are  
XX useful in the diagnosis, treatment and prevention of various types of  
XX disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
XX chickens or sheep. A pathological condition can be determined by  
XX detecting the presence or absence of a mutation in a cardiovascular  
XX system antigen polynucleotide. The treatable disorders include autoimmune  
XX diseases such as rheumatoid arthritis, hyperproliferative disorders such  
XX as neoplasms of the breast or liver, cardiovascular disorders such as  
XX cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,  
XX nervous system disorders such as Alzheimer's disease, infections caused  
XX by bacteria, viruses and fungi, ocular disorders such as corneal  
XX infection, endocrine disorders such as premature labour and infertility,  
XX gastrointestinal disorders such as Crohn's disease, renal disorders such  
XX as glomerulonephritis and respiratory disorders such as asthma and  
XX pleurisy. The polypeptides can also be used to aid wound healing, to  
XX prevent skin aging due to sunburn, to maintain organs before  
XX transplantation, to regenerate tissues and in chemotaxis.  
XX Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
Query Match 83.5%; Score 42.6; DB 22; Length 4962;  
Best Local Similarity 88.2%; Pred. No. 1.9e-06;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCSGGTCTCTACAAAAATACA 51  
Db 2690 TCAAGACCAGCCTGGACACATGGTGAACCCGTGTCTCTACAAAAATACA 2740  
RESULT 15  
AAV17618/C  
ID AAV17618 standard; DNA; 12047 BP.  
XX  
XX AAV17618;  
XX  
XX 21-JUL-1998 (first entry)  
XX  
XX Homo sapiens glandular kallikrein-1 promoter region.  
XX  
XX Promoter; enhancer; 5' non-coding region; hKLK2; kallikrein; hGK-1;  
XX PSA; prostate specific antigen; prostate cancer; treatment; ss.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX enhancer 6859..8627  
XX /\*tag= a  
XX /note= "enhancer active fragment"  
XX promoter 11420..12047  
XX /\*tag= b  
XX /note= "promoter active fragment"  
XX mRNA 12013..12047  
XX /\*tag= c  
XX /note= "transcription initiation at 12013"  
XX  
XX WO9805797-A1.  
XX  
XX 12-FEB-1998.  
XX  
XX 06-AUG-1997; 97WO-US13888.  
XX  
XX 04-AUG-1997; 97US-0692759.  
XX 06-AUG-1996; 96US-0692759.  
XX  
XX (CALY-) CALYDON.  
XX  
XX Henderson DR, Lamparski HG, Schuur ER, Yu DC;  
XX WPI; 1998-145628/13.  
XX  
XX Screening drugs for treatment of prostate cancer - uses prostate  
XX specific antigen expressing cells  
XX  
XX Disclosure; Pages 25-28; 44pp; English.  
XX  
XX The sequence is that of the non-coding region of human glandular  
XX kallikrein-1 (hKLK2) which contains promoter and enhancer active  
XX regions. This transcription initiation regulatory region can be used  
XX as part of a method for screening drugs for the treatment of prostate  
XX cancer employing prostate specific antigen (PSA) expressing cells  
XX comprising an expression construct which comprises a transcriptional  
XX initiation region of the PSA enhancer, a promoter and a gene whose  
XX expression product provides a detectable signal, where the gene is  
XX under the transcriptional control of the initiation region, comprises:  
XX (a) combining the PSA expressing cells with a candidate drug in the  
XX presence of an androgen for sufficient time for detectable expression  
XX of the gene, and (b) detecting the level of expression of the gene as  
XX compared to the level of expression in the absence of the candidate  
XX drug. The method can be used to evaluate the potential of a



CC compound as a therapeutic agent for the treatment of prostate  
CC cancer. Particularly, anti-androgenic activity can be evaluated as  
CC indicative of therapeutic effects for prostate cancer, although any  
CC compound which modifies the expression of a prostate-specific gene  
CC may be considered a candidate compound.

XX

SQ Sequence 12047 BP; 3051 A; 2484 C; 2869 G; 3643 T; 0 other;

Query Match 83.5%; Score 42.6; DB 19; Length 12047;

Best Local Similarity 88.2%; Pred. No. 2.4e-06;

Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGACAACCTTGGAGAACCGGTCTCTACAAAAATACA 51

|||||

Db 1985 TCAAGACCGCTGGCCACATGGCAAAACCGCGTCTCTACAAAAATACA 1935

Search completed: February 11, 2003, 02:04:45  
Job time : 24.723 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 21:06:50 ; Search time 141.088 Seconds  
(without alignments)  
8457.264 Million cell updates/sec

Title: us-09-942-310-2\_copy\_175\_215  
Perfect score: 41  
Sequence: 1 cctatctctactgaaatay.....aaaagctagacgtggtggca 41

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_cm:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pi:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	40.6	99.0	1669	6	AX207224	Sequence
2	40.6	99.0	1680	6	AX394457	Sequence
3	40.6	99.0	9432	6	AX394456	Sequence
4	40.6	99.0	9432	9	HUMCYP2D6	Human CYP2C
5	40.6	99.0	13677	9	HSCYFP2D7B	Human CYP2D
c	34.2	83.4	170346	9	AC020626	Homo sapi
7	34.2	83.4	172611	9	AC019129	Homo sapi
8	34.2	83.4	187486	2	AC079840	Homo sapi
9	34.2	83.4	207266	2	AC127379	Homo sapi
10	33.2	81.0	170307	9	CNS00002	Human chr
11	33.2	81.0	210044	9	AC016586	Homo sapi
12	32.6	79.5	39528	9	AL591437	Human DNA
c	32.6	79.5	103606	9	AC005004	Homo sapi
13	32.6	79.5	110000	2	HSY313F4_0	Homo sapi
14	32.6	79.5	112088	9	AC005216	Homo sapi
15	32.6	79.5	117928	9	AC026413	Homo sapi
c	32.6	79.5	124289	9	AC010285	Homo sapi
17	32.6	79.5	126503	9	AC008114	Homo sapi
18	32.6	79.5	129564	9	AC087302	Homo sapi
19	32.6	79.5	142552	9	HS591B8	Human DNA
c	32.6	79.5	146328	9	HSY738F9	Human DNA
21	32.6	79.5	153376	9	AC027026	Homo sapi
22	32.6	79.5	155952	9	AC004534	Homo sapi
23	32.6	79.5	160530	2	AC025631	Homo sapi
24	32.6	79.5	161151	2	AP000876	Homo sapi
25	32.6	79.5	163937	2	AC091887	Homo sapi
26	32.6	79.5	165242	9	AC009298	Homo sapi
c	32.6	79.5	167932	2	AC068133	Homo sapi
28	32.6	79.5	168071	2	AL391479	Homo sapi
29	32.6	79.5	168324	9	AL133391	Human DNA
30	32.6	79.5	168753	9	AC006120	Homo sapi
c	32.6	79.5	170078	9	AC015914	Homo sapi
32	32.6	79.5	172309	9	AC018553	Homo sapi
c	32.6	79.5	173071	9	AC009477	Homo sapi
34	32.6	79.5	175511	2	AC033924	Homo sapi
35	32.6	79.5	179610	2	AC124892	Homo sapi
36	32.6	79.5	182105	2	AC036149	Homo sapi
37	32.6	79.5	184092	2	AC124858	Homo sapi
38	32.6	79.5	187628	9	AC011311	Homo sapi
c	32.6	79.5	188800	2	AC015846	Homo sapi
40	32.6	79.5	189013	9	AC025277	Homo sapi
41	32.6	79.5	193267	2	AC009409	Homo sapi
c	32.6	79.5	200823	9	AC009652	Homo sapi
43	32.6	79.5	211945	9	AC006238	Homo sapi
c	32.6	79.5	212730	2	AL354975	Homo sapi
45	32.6	79.5	212730	2	AL354975	Homo sapi

ALIGNMENTS

RESULT 1	AX207224	AX207224	Sequence	1669 bp	DNA	linear	PAT 30-AUG-2001
AX207224	LOCUS	Sequence	1 from Patent WO0155432.				
DEFINITION	AX207224	AX207224	Sequence				
ACCESSION	AX207224	AX207224	Sequence				
VERSION	AX207224.1	GI:15394976	GI:15394976				
KEYWORDS			synthetic construct.				
SOURCE			synthetic construct				
ORGANISM			artificial sequences.				
REFERENCE			1 (bases 1 to 1669)				
AUTHORS			Raimundo, S. and Zanger, U.				
TITLE			Polymorphisms in the human cyp2d6 gene promoter region and their use in diagnostic and therapeutic applications				
JOURNAL			Patent: WO 0155432-A 1 02-AUG-2001;				

```

FEATURES
  source
    Epidauros Biotechnologie AG (DE)
    Location/Qualifiers
      1..1669
      /organism="synthetic construct"
      /db_xref="taxon:32630"
      /note="artificial sequence"
BASE COUNT
  413 a 376 c 534 g 338 t 8 others
ORIGIN
  Query Match
  Best Local Similarity 99.0%; Score 40.6; DB 6; Length 1659;
  Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 41
  Db 175 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 215
RESULT 2
AX394457
LOCUS
  AX394457 1680 bp DNA linear PAT 18-MAY-2002
DEFINITION
  Sequence 2 from Patent W00218638.
ACCESSION
  AX394457
VERSION
  AX394457.1 GI:21065595
KEYWORDS
  human.
SOURCE
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1
  AUTHORS
    Risinger,C., Andersson,M.K., Lewander,T. and Ollasson,E.
  TITLE
    Detection of cyp2d6 polymorphisms
  JOURNAL
    Patent: WO 0218638-A 2 07-MAR-2002;
    Gemini Genomics PLC (GB)
FEATURES
  source
    Location/Qualifiers
      1..1680
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
BASE COUNT
  413 a 379 c 539 g 342 t 7 others
ORIGIN
  Query Match
  Best Local Similarity 99.0%; Score 40.6; DB 6; Length 1680;
  Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 41
  Db 175 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 215
RESULT 3
AX394456
LOCUS
  AX394456 9432 bp DNA linear PAT 18-MAY-2002
DEFINITION
  Sequence 1 from Patent W00218638.
ACCESSION
  AX394456
VERSION
  AX394456.1 GI:21065594
KEYWORDS
  human.
SOURCE
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1
  AUTHORS
    Risinger,C., Andersson,M.K., Lewander,T. and Ollasson,E.
  TITLE
    Detection of cyp2d6 polymorphisms
  JOURNAL
    Patent: WO 0218638-A 1 07-MAR-2002;
    Gemini Genomics PLC (GB)
FEATURES
  source
    Location/Qualifiers
      1..9432
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
BASE COUNT
  1964 a 2647 c 2976 g 1845 t
ORIGIN

```

```

Query Match
Best Local Similarity 99.0%; Score 40.6; DB 6; Length 9432;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 41
Db 175 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 215
RESULT 4
HUMCYP2D6
LOCUS
  HUMCYP2D6 9432 bp DNA linear PRI 22-NOV-1994
DEFINITION
  Human cytochrome P450 IID6 (CYP2D6) gene, complete cds.
ACCESSION
  M33388
VERSION
  M33388.1 GI:181303
KEYWORDS
  cytochrome P450; cytochrome P450 IID6.
SOURCE
  Human DNA, clone lambda2D-18/2.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 9432)
  AUTHORS
    Kimura,S., Umeno,M., Skoda,R.C., Meyer,U.A. and Gonzalez,F.J.
  TITLE
    The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and
    identification of the polymorphic CYP2D6 gene, a related gene, and
    a pseudogene
  JOURNAL
    Am. J. Hum. Genet. 45 (6), 889-904 (1989)
  MEDLINE
    90072069
  PUBMED
    2574001
COMMENT
  Draft entry and computer-readable sequence for {Am. J. Hum. Genet.
  45, 889-904 (1989)} kindly submitted
  by S.Kimura, 29-MAR-1990.
FEATURES
  source
    Location/Qualifiers
      1..9432
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /map="22q13.1"
      join(1532..1799,2503..2674,3225..3377,3466..3626,
      4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)
      /gene="CYP2D6"
      join(1532..1799,2503..2674,3225..3377,3466..3626,
      4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)
      /gene="CYP2D6"
      /product="cytochrome P450 IID6"
      /note="G00-132-127"
      1532..1799
      /partial
      /gene="CYP2D6"
      /note="cytochrome P450 IID6; G00-132-127"
      /number=1
      join(1620..1799,2503..2674,3225..3377,3466..3626,
      4060..4236,4427..4568,4776..4963,5418..5559,5658..5836)
      /gene="CYP2D6"
      /codon_start=1
      /product="cytochrome P450 IID6"
      /protein_id="AAA53500.1"
      /db_xref="GI:181304"
      /db_xref="GDB:G00-132-127"
      /translation="MGLEALVPLAVIVAFLLVLDLMHRRQRWAARYPPGPLPLPLGLG
      NLLHVDVQNTPEYCFDQLRRRGDFVSLQLANTPVVVLNGLAAVREALVTHGEDFADRP
      CAVPTIQLGFRSQGVFLARYGPAWREQRFSVSTLRNLGKGLKLEQWTEEAACL
      PAFANHSGRFPRNGLLDKAVSNVIAISLTCGRRFYDDPRFLRLDLAQGLKEESG
      FLREVLNAPVLLHPIPALAGKVLRFKAFLTQLDELTEHRTWDPAQPPRDLTEAFL
      AEMEKAKGNPESFNDELRIIVADLFSAQVMTTSTTLCMLLILHPPDVQRVQOE
      IDDVIGQVRRPEMGDAQHPYTTAVIHEVOPEGDIVPLGYVTHMTSRDIEVQGFRLPGK
      TTLTLNLSVLKDEAVWEKPRFPHFELDQGHFVKAPEAFLPFSAGRRACLGLPELAR
      MELFLFTSLVLQKHEFSF SVPTGQPRPSHHGVFAFLVSPSPVELCAVPR"
      1800..2502
      /gene="CYP2D6"
      /note="G00-132-127"
      /number=1
      2503..2674
      /gene="CYP2D6"
      /note="G00-132-127"
  intron
  exon

```



```

intron      4541.. 4734
            /gene="CYP2D7BP"
            /number=6
exon        4735.. 4922
            /gene="CYP2D7BP"
            /number=7
intron      4923.. 5376
            /gene="CYP2D7BP"
            /number=7
exon        5377.. 5518
            /gene="CYP2D7BP"
            /number=8
intron      5519.. 5616
            /gene="CYP2D7BP"
            /number=8
exon        5617.. 5868
            /gene="CYP2D7BP"
            /number=9
repeat_region 8267.. 8306
            /note="ACCCCTCCCC"

BASE COUNT 3066 a 3775 c 4107 g 2729 t

Query Match          99.08;   Score 40.6;   DB 9;   Length 13677;
Best Local Similarity 97.68;   Pred. No. 0.0002;
Matches 40;   Conservative 1;   Mismatches 0;   Indels 0;   Gaps 0;

Oy 1 CCTATCTCTACTGAAATATAYAAAAAGCTAGACGTGGTGGCA 41
      |||||||
Db 175 CCTATCTCTACTGAAATATAYAAAAAGCTAGACGTGGTGGCA 215

RESULT 6
AC020626/c
LOCUS      AC020626               170345 bp    DNA    linear    PRI 28-APR-2000
DEFINITION Homo sapiens 3 BAC Rpl1-41F5 (Roswell Park Cancer Institute Human
ACCESSION  AC020626
VERSION     AC020626.6  GI:76566676
KEYWORDS   HTG.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 170346)
AUTHORS   Muzny,D.M., Adams,C., Bailey,M., Barberia,J., Blankenburg,K.,
            Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
            Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
            David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,
            Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
            Forcum-Tansey,J., Frantz,P., Ganesh,R., Garcia,D.K., Gorrell,J.H.,
            Gorrell,L.L., Guevara,W., Harris,K., He,X., Hernandez,J.,
            Hodgson,A., Hogue,M., Holloway,C., Hosak,H., Jackson,L.E.,
            Jackson,L., Jia,Y., Jones,M., Kelly,S., Kondejewski,N., Kong,Y.,
            Kovar,C., Leal,B., Li,Z., Lichtarge,O., Liu,J., Liu,W., Logan,O.,
            Lozado,R.J., Lu,J., Lucier,R., Martin,R., Martinez,C., McLeod,M.P.,
            Mei,G., Moore,S., Moorish,T., Morgan,M., Morris,S., Nash,S.,
            Nelson,A., Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B.,
            Paxton,S., Payton,B., Perez,L., Pu,L.L., Quiles,M., Reiter,D.,
            Rives,M., Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H.,
            Simon,M., Sparks,A., Stamps,A., Suggang,R., Tabor,P., Taylor,T.,
            Vasquez,L., Vinson,R., Vo,Q., Wabhan,M., Watlington,S.,
            Weinstock,G., Weinstock,I.R., Williamson,A., Worley,K., Wren,J.,
            Wrensford,G., Yu,W., Zhou,X., Naylor,S.L., Nelson,D. and Gibbs,R.
Direct Submission
Unpublished
            2 (bases 1 to 170346)
            Worley,K.C.
Direct Submission
Submitted (07-JAN-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE  3 (bases 1 to 170346)
            Worley,K.C.
TITLE     Worley,K.C.
JOURNAL   Submitted (27-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE  4 (bases 1 to 170346)
            Worley,K.C.
TITLE     Worley,K.C.
JOURNAL   Submitted (28-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT    On Apr 27, 2000 this sequence version replaced gi:7331305.
            INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
            gc-helpebcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STSS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL: http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.
html.

QUALSTAT-REPORT-----
----- Summary Statistics -----
Contig length: 170346
Phrap values in estimate: 168979
Average error rate (BCM-Phrap estimate): 0.000470003
Fraction of Phrap values less than 40 : 0.049012
Number of consensus changing edits: 32
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context
1965 caaaggaaga(n)gaagttaagt caaaggaaga(t)gaagttaagt
3255 tttttgaga(n)ggagttttgc tttttgaga(t)ggagttttgc
28741 tatctggtc(t)acatggtgaa tatctggtc(a)acatggtgaa
28752 acatggtgaa(n)ccctgtctct acatggtgaa(a)ccctgtctct
34551 caccctggca(n)aaatggaata caccctggca(a)aaatggaata
42694 aacaggaacc(n)ccgtgaaac aacaggaacc(a)ccgtgaaac
43319 catgagaat(n)ttctgaacn catgagaat(c)ctctgaacn
43320 catgagaat(n)ttctgaacn catgagaat(c)ctctgaacn
43329 ncttgaacc(n)ngagggcgaa gcttgaacc(t)ggagggcgaa
43330 tcttgaacc(n)ggagggcgaa gcttgaacc(t)ggagggcgaa

```



NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute vector: pBACE3.6 (<http://daccpac.med.buffalo.edu>)

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC087053; the clone sequenced to the right is RP11-2415, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-559M23; actual end is at base position 17064 of RP11-2415.

Single plasmid coverage exists between bases 30202 to 30267.

FEATURES	Source	Location/Qualifiers
		1..172611
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="2"
		/map="2"
		/clone="RP11-559M23"
		/clone_lib="RPCI-11"
repeat_region		1..1439
		/rpt_family="ERV1"
repeat_region		1440..1738
		/rpt_family="Alu"
repeat_region		1739..1964
		/rpt_family="ERV1"
repeat_region		1977..2082
		/rpt_family="ERV1"
repeat_region		2087..2218
		/rpt_family="Alu"
repeat_region		3209..3259
		/rpt_family="TTTTGn"
repeat_region		3230..3551
		/rpt_family="Alu"
repeat_region		3406..3426
		/rpt_family="AT-rich"
repeat_region		3554..3580
		/rpt_family="AT-rich"
repeat_region		4188..4453
		/rpt_family="Alu"
repeat_region		4453..4634
		/rpt_family="GA-rich"
repeat_region		4672..4976
		/rpt_family="Alu"
repeat_region		4953..4976
		/rpt_family="(A)n"
repeat_region		5162..5399
		/rpt_family="L2"
repeat_region		5430..5588
		/rpt_family="MER1_type?"
		5559..5594
		/rpt_family="AT-rich"
		5608..5628
		/rpt_family="TTTTA)n"
		5611..5744
		/rpt_family="Alu"
		6207..6265
		/rpt_family="MER1_type?"
		6338..6625
		/rpt_family="MER1_type?"
		6543..6579
		/rpt_family="AT-rich"
		6671..6732
		/rpt_family="Alu"
		6755..7031
		/rpt_family="ERV1"
		7059..7098
		/rpt_family="ERV1"
		7096..7166
		/rpt_family="MER1_type?"
		7428..7712
		/rpt_family="Alu"
		7720..8056
		/rpt_family="L2"
		8429..8537
		/rpt_family="Alu"
		8708..8756
		/rpt_family="L2"
		9394..9432
		/rpt_family="TTTTA)n"
		9417..9708
		/rpt_family="Alu"
		9710..9730
		/rpt_family="TTTTA)n"
		9712..10012
		/rpt_family="Alu"
		10226..10618
		/rpt_family="ERV1"
		11969..12201
		/rpt_family="MIR"
		12637..12795
		/rpt_family="MIR"
		12930..13232
		/rpt_family="Alu"
		13351..13421
		/rpt_family="MER103"
		13422..13717
		/rpt_family="Alu"
		13718..13756
		/rpt_family="MER103"
		13831..14034
		/rpt_family="MIR"
		14518..14566
		/rpt_family="T-rich"
		15603..15624
		/rpt_family="AT-rich"
		16436..16680
		/rpt_family="L2"
		16672..16710
		/rpt_family="(T)n"
		16681..16978
		/rpt_family="Alu"
		16979..17051
		/rpt_family="L2"
		17123..17255
		/rpt_family="Alu"
		17238..17264
		/rpt_family="AT-rich"
		17347..17732
		/rpt_family="L2"
		17767..17864
		/rpt_family="MER53"





\* 167567 170676: contig of 3110 bp in length  
 \* 170677 170776: gap of unknown length  
 \* 170777 174692: contig of 3916 bp in length  
 \* 174693 174792: gap of unknown length  
 \* 174793 177876: contig of 3084 bp in length  
 \* 177877 177976: gap of unknown length  
 \* 177977 180949: contig of 2973 bp in length  
 \* 180950 181049: gap of unknown length  
 \* 181050 184627: contig of 3578 bp in length  
 \* 184628 184727: gap of unknown length  
 \* 184728 187486: contig of 2759 bp in length.

# FEATURES

source  
 1. .187486  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="13"  
 /clone="RP11-316J16"  
 BASE COUNT 52952 a 39656 c 39493 g 52710 t 2675 others  
 ORIGIN

Query Match 83.4%; Score 34.2; DB 2; Length 187486;  
 Best Local Similarity 87.8%; Pred. No. 0.033;  
 Matches 36; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAATAATAYAAAAAGCTAGACGTGGTGGA 41  
 Db 349 CCTGTCTCTACTAAAAATACAAAAACCTAGGCGTGGTGGA 389

RESULT 9  
 AC127379 207266 bp DNA linear HTG 13-AUG-2002  
 LOCUS  
 DEFINITION Homo sapiens chromosome 2 clone RP11-808E20, WORKING DRAFT  
 AC127379  
 AC127379.2 GI:22213495  
 VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
 KEYWORDS human.  
 SOURCE Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 AUTHORS Waterston,R.H.  
 TITLE The sequence of Homo sapiens clone  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 207266)  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-JUL-2002) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 3 (bases 1 to 207266)  
 AUTHORS Waterston,R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-AUG-2002) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA  
 COMMENT On Aug 13, 2002 this sequence version replaced gi:21747801.

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc/index.shtml>  
 Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)  
 ----- Project Information -----  
 Center project name: H.NH0808E20  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 0%  
 Sequencing vector: plasmid; 100%  
 Chemistry: Dye-primer ET; 0% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319

Consensus quality: 206829 bases at least Q40  
 Consensus quality: 207045 bases at least Q30  
 Consensus quality: 207263 bases at least Q20

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 1 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 207266: contig of 207266 bp in length.

# FEATURES

source  
 1. .207266  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /clone="RP11-808E20"

misc\_feature  
 1. .207266  
 /note="assembly\_name:Contig8"  
 BASE COUNT 64149 a 42249 c 42430 g 58438 t  
 ORIGIN

Query Match 83.4%; Score 34.2; DB 2; Length 207266;  
 Best Local Similarity 87.8%; Pred. No. 0.032;  
 Matches 36; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAATAATAYAAAAAGCTAGCGTGGTGGA 41  
 Db 203378 CCCATCTCTACTAAAAATACAAAAACCCAGATGTGGTGGA 203418

RESULT 10  
 CNS00002  
 LOCUS  
 DEFINITION

Human chromosome 14 DNA sequence BAC R-96617 of library RPCI-11  
 from chromosome 14 of Homo sapiens (Human), complete sequence.

ACCESSION AL049777  
 VERSION AL049777.5 GI:7838309  
 KEYWORDS HTG.  
 SOURCE human.

# ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 170307)  
 AUTHORS Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,  
 Brottier,P., Cattolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,  
 Levy,M., Eckenberg,R., Bruls,T., Deberardinis,V., Cruaud,C.,  
 Gyapay,G., Saurin,W. and Weissbach,J.  
 TITLE Sequencing of the human chromosome 14  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 170307)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :  
 BP 191 91006 EVRY cedex - FRANCE (E-mail : [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr))  
 COMMENT On May 15, 2000 this sequence version replaced gi:6468471.

----- Genome Center  
 Center: Genoscope / Centre National de Sequencage  
 Center code: GS  
 Web site: <http://www.genoscope.cns.fr/>  
 Contact: [SeqRef@genoscope.cns.fr](mailto:SeqRef@genoscope.cns.fr)

# COMMENT

The following BAC sequence is oriented from the T7 to the SP6 end.  
 Upstream BAC (overlapping the T7 end) : R-412H8  
 Downstream BAC (overlapping the SP6 end) : R-30H9 (AC=ALL22126)  
 ----- Summary Statistics -----  
 Assembly program: Phrap; version 1.0  
 Quality coverage: 5.88x in Q20 bases; sum-of-contigs

## Overall quality chart :

```

Range : bases
0      :
1 - 9  : 7
10 - 19 : 36
20 - 29 : 156
30 - 39 : 585
40 - 49 : 1587
50 - 59 : 6145
60 - 69 : 21850
70 - 79 : 56146
80 - 89 : 52135
90 - 99 : 31660

```

Percentage of bases with a quality value >= 40 : 99 %.

```

FEATURES
source
1. .170307
   Location/Qualifiers
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="14"
   /clone="R-96617"
   /clone_lib="RPCr-11"
   43803..43927
   /note="matching EMBL:H23232
dbSTS:STS26386
RHdb:RH53834
Identified using the e-PCR software (G. Schuler)"
94717..95047
   /note="matching EMBL:X74142
RHdb:RH14114
dbSTS:STS24699
Identified using the e-PCR software (G. Schuler)"
95767..95916
   /note="matching EMBL:D59942
RHdb:RH53545
dbSTS:STS7285
Identified using the e-PCR software (G. Schuler)"
119069..119269
   /note="matching EMBL:242192
RHdb:RH25525
dbSTS:STS13977
Identified using the e-PCR software (G. Schuler)"
120170..120301
   /note="matching EMBL:G32636
RHdb:RH67328
dbSTS:STS47255
Identified using the e-PCR software (G. Schuler)"
120257..120433
   /note="matching EMBL:238448
RHdb:RH53471
dbSTS:STS16988
Identified using the e-PCR software (G. Schuler)"
54312 a 31571 c 31363 g 53061 t

Query Match      81.0%; Score 33.2; DB 9; Length 170307;
Best Local Similarity 87.5%; Pred. No. 0.075;
Matches 35; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATATAYAAAAAGCTAGACGTGGTGGC 40
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37574 CCTGCTCTACTATAAATACAAAAAGCCGCGTGGTGGC 37613

RESULT 11
AC016586      210044 bp DNA linear PRI 31-JUL-2002
LOCUS      Homo sapiens chromosome 19 clone CTD-2622113, complete sequence.
DEFINITION
AC016586
VERSION
AC016586.8 GI:22024554
KEYWORDS
HTG.
SOURCE
human.
ORGANISM
Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 210044)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished
2 (bases 1 to 210044)
DOE Joint Genome Institute.
Direct Submission
Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 210044)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (31-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jul 31, 2002 this sequence version replaced gi:15022014.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.9% of Sequence;
Estimated Total Number of Errors is 0.3.
NOTE: Shatter libraries failed to resolve repeat region
65151-65432. Unsure number of repeat copies 65151-65432. Forced
Join 65401.
Location/Qualifiers
1. .210044
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="19"
   /clone="CTD-2622113"
   65151..65432
   /note="NOTE: Shatter libraries failed to resolve repeat
region 65151-65432. Unsure number of repeat copies
65151-65432. Forced Join 65401."
BASE COUNT 46512 a 59070 c 58070 g 46392 t
ORIGIN

Query Match      81.0%; Score 33.2; DB 9; Length 210044;
Best Local Similarity 87.5%; Pred. No. 0.074;
Matches 35; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATATAYAAAAAGCTAGACGTGGTGGC 40
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 69421 CCTATCTCTACTATAAATACAAAAAGCTGGCGTGGTGGC 69460

RESULT 12
AL591437      39528 bp DNA linear PRI 17-NOV-2001
LOCUS      Human DNA sequence from clone RP13-223M5 on chromosome X, complete
sequence.
DEFINITION
AL591437
VERSION
AL591437.6 GI:17017770
KEYWORDS
HTG.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 39528)
Direct Submission
Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonequest@sanger.ac.uk
On Nov 20, 2001 this sequence version replaced qi:14346133.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

```

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/ChrX> RP13-223M5 is from the library RPCI-13.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP13-223M5 It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true left end of clone RP13-192B19 is at 37529 in this sequence. The true right end of clone RP1-314H24 is at 2000 in this sequence.

```

FEATURES             Location/Qualifiers
     source            1..39528
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="X"
                     /clone="RP13-223M5"
                     /clone_lib="RPCI-13.1"
BASE COUNT          12293 a 7301 c 7036 g 12898 t
ORIGIN
Query Match          79.5%; Score 32.6; DB 9; Length 39528;
Best Local Similarity 85.4%; Pred. No. 0.13;
Matches 35; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CCTATCTCTACTGAAATAYAAAAAGCTAGCGTGGCGCA 41
    ||| ||||| ||||| ||||| ||| ||||| |||||
Db 30618 CCTGCTCTACTAAATACAAAAAATAGCGTGGCGCA 30658

RESULT 13
AC005004/c
LOCUS                103606 bp DNA linear PRI 21-DEC-1999
DEFINITION            Homo sapiens PAC clone RP3-403E2 from 22, complete sequence.
ACCESSION              AC005004
VERSION                AC005004.3 GI:66241123
KEYWORDS
SOURCE
ORGANISM              Homo sapiens
                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE              1 (bases 1 to 103606)
AUTHORS                Waterston,R.
TITLE                  The sequence of Homo sapiens PAC clone RP3-403E2
JOURNAL                Unpublished
REFERENCE              2 (bases 1 to 103606)
AUTHORS                Waterston,R.H.
TITLE                  Direct Submission
JOURNAL                Submitted (12-JUN-1998) Genome Sequencing Center, Washington
                     University School of Medicine, 4444 Forest Park Parkway, St. Louis,
                     MO 63108, USA
REFERENCE              3 (bases 1 to 103606)
AUTHORS                Waterston,R.
TITLE                  Direct Submission
JOURNAL                Submitted (14-JAN-1999) Department of Genetics, Washington
                     University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE              4 (bases 1 to 103606)
AUTHORS                Waterston,R.

```

```

TITLE
JOURNAL                Direct Submission
                     Submitted (20-JAN-1999) Department of Genetics, Washington
                     University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE              5 (bases 1 to 103606)
AUTHORS                Waterston,R.
TITLE                  Direct Submission
JOURNAL                Submitted (11-DEC-1999) Department of Genetics, Washington
                     University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE              6 (bases 1 to 103606)
AUTHORS                Waterston,R.
TITLE                  Direct Submission
JOURNAL                Submitted (14-DEC-1999) Department of Genetics, Washington
                     University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE              7 (bases 1 to 103606)
AUTHORS                Waterston,R.
TITLE                  Direct Submission
JOURNAL                Submitted (21-DEC-1999) Department of Genetics, Washington
                     University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT               On Dec 21, 1999 this sequence version replaced gi:5560935.
                     ----- Genome Center
                     Center: Washington University Genome Sequencing Center
                     Center code: WUGSC
                     Web site: http://genome.wustl.edu/gsc
                     Contact: sapiens@watson.wustl.edu
                     ----- Summary Statistics
                     -----
                     Center project name: H_DJ403E2
                     -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 22 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR22>, send [mailto:egreen@nhgri.nih.gov](mailto:mailto:egreen@nhgri.nih.gov), or see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-3, prepared by Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>) using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.

The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong.

#### VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:

The actual start of this clone is at base position 1 of RP3-403E2; actual end is at 103606 of RP3-403E2.

```

FEATURES             Location/Qualifiers
     source            1..103606
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="22"
                     /clone="RP3-403E2"
                     /clone_lib="RPCI-3"
                     1..235
                     /rpt_family="Alu"
repeat_region

```

```
repeat_region 282. .591 /rpt_family="Alu"
repeat_region 697. .719 /rpt_family="Alu"
repeat_region 903. .1198 /rpt_family="(TAAAA)n"
repeat_region 1381. .1689 /rpt_family="Alu"
repeat_region 1834. .2139 /rpt_family="Alu"
repeat_region 2165. .2331 /rpt_family="Alu"
repeat_region 2332. .2635 /rpt_family="Alu"
repeat_region 2636. .2755 /rpt_family="Alu"
repeat_region 2758. .3034 /rpt_family="Alu"
repeat_region 3047. .3353 /rpt_family="Alu"
repeat_region 3355. .3444 /rpt_family="Alu"
repeat_region 3467. .3754 /rpt_family="Alu"
repeat_region 3890. .3923 /rpt_family="AT_rich"
repeat_region 3970. .4266 /rpt_family="Alu"
repeat_region 4705. .4985 /rpt_family="Alu"
repeat_region 4986. .5012 /rpt_family="Alu"
repeat_region 5024. .5316 /rpt_family="(TAAA)n"
misc_feature 5596. .5768 /note="match to EST H26488 (NID:9896479)"
repeat_region 11261. .11291 /rpt_family="(TAAA)n"
repeat_region 11293. .11439 /rpt_family="Alu"
repeat_region 11440. .11740 /rpt_family="Alu"
repeat_region 11741. .11764 /rpt_family="Alu"
repeat_region 11765. .12066 /rpt_family="Alu"
repeat_region 12067. .12154 /rpt_family="Alu"
repeat_region 12155. .12196 /rpt_family="Retroviral"
repeat_region 12230. .12252 /rpt_family="AT_rich"
repeat_region 12305. .12588 /rpt_family="Alu"
repeat_region 12589. .12622 /rpt_family="(CAAAA)n"
repeat_region 12642. .12956 /rpt_family="Alu"
repeat_region 12957. .13094 /rpt_family="L2"
repeat_region 13197. .13252 /rpt_family="MIR"
repeat_region 13253. .13547 /rpt_family="Alu"
repeat_region 13548. .13693 /rpt_family="MIR"
repeat_region 13964. .14275 /rpt_family="Alu"
repeat_region 14866. .15091 /rpt_family="Alu"
repeat_region 15150. .15348 /rpt_family="L1"
repeat_region 15661. .15955
```

```
repeat_region /rpt_family="Alu"
16038. .16344 /rpt_family="Alu"
17368. .17892 /rpt_family="MERL_type"
17928. .18210 /rpt_family="Alu"
18628. .18870 /rpt_family="Alu"
19178. .19196 /rpt_family="AT_rich"
19197. .19481 /rpt_family="Alu"
19482. .19498 /rpt_family="AT_rich"
19673. .19698 /rpt_family="Alu"
19703. .19943 /rpt_family="Alu"
19947. .20079 /rpt_family="Alu"
20745. .21047 /rpt_family="Alu"
21213. .21516 /rpt_family="Alu"
21583. .21696 /rpt_family="MaLR"
21844. .22137

Query Match 79.5%; Score 32.6; DB 9; Length 103606;
Best Local Similarity 85.4%; Pred. No. 0.13;
Matches 35; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTACACGTGGTGCCA 41
Db 75134 CCCATCTCTACTAAATAATAAAAAAGCCAGGTGTGGTGCCA 75094

RESULT 14
HSY313F4_0
WPCOMMENT
Sequence split into 4 fragments LOCUS HSY313F4 Accession AL023808
Fragment Name Begin End
HSY313F4_0 1 110000
HSY313F4_1 100001 210000
HSY313F4_2 200001 310000
HSY313F4_3 300001 368509
LOCUS HSY313F4 368509 bp DNA linear HTG 02-FEB-2002
DEFINITION Homo sapiens chromosome 10 clone XX-Y313F4, *** SEQUENCING IN
PROGRESS ***, 18 unordered pieces.
ACCESSION AL023808
VERSION 4 GI:18494913
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS McMurray,A.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 4, 2002 this sequence version replaced gi:9801261.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: Y313F4
----- Summary Statistics
Assembly program: XGAP4; version 4.5
```



JOURNAL REFERENCE AUTHORS	Comparison Analysis (SCAN) System	
	Unpublished	
	3 (bases 1 to 112088)	
TITLE JOURNAL	Kimberly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T., Pittluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.	
	Direct Submission	
	Submitted (01-JUL-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A.	
COMMENT	Sequence submitted by:	
FEATURES	DOE Joint Genome Institute.	
	Location/Qualifiers	
Source	1. .112088	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="5"	
	/map="5q"	
	/clone="BAC 282B7"	
	278. .400	
misc_feature	/note="GRAIL 2 excellent exon, frame 1"	
repeat_region	2373. .2655	
repeat_region	/rpt_family="Alu"	
repeat_region	complement(6099. .8811)	
repeat_region	/rpt_family="L1"	
repeat_region	10784. .10913	
repeat_region	/rpt_family="MER45"	
repeat_region	12576. .12619	
	/note="(ATGG)11"	
	/rpt_type=tandem	
repeat_region	/rpt_unit=ATGG	
repeat_region	complement(14548. .16322)	
repeat_region	/rpt_family="L1"	
repeat_region	complement(18472. .18652)	
repeat_region	/rpt_family="Alu"	
repeat_region	18788. .19074	
repeat_region	/rpt_family="Alu"	
repeat_region	19089. .19122	
	/note="(AC)17"	
	/rpt_type=tandem	
repeat_region	/rpt_unit=AC	
repeat_region	complement(19399. .19697)	
repeat_region	/rpt_family="Alu"	
repeat_region	complement(20481. .20950)	
repeat_region	/rpt_family="L1"	
repeat_region	22980. .23163	
repeat_region	/rpt_family="MIR"	
repeat_region	24364. .24411	
	/note="(AC)24"	
	/rpt_type=tandem	
repeat_region	/rpt_unit=AC	
repeat_region	complement(24681. .24880)	
repeat_region	/rpt_family="MER20"	
repeat_region	complement(25387. .25503)	
repeat_region	/rpt_family="MIR"	
repeat_region	29226. .29369	
repeat_region	/rpt_family="MIR"	
repeat_region	29654. .29690	
	/note="(T)37"	
	/rpt_type=tandem	
repeat_region	/rpt_unit=T	
repeat_region	complement(29673. .29961)	
repeat_region	/rpt_family="Alu"	
repeat_region	31039. .31321	
repeat_region	/rpt_family="Alu"	
repeat_region	32487. .32904	
repeat_region	/rpt_family="Alu"	
repeat_region	complement(33328. .33454)	
repeat_region	/rpt_family="MIR"	
repeat_region	35276. .35369	
repeat_region	/rpt_family="L1"	
repeat_region	35617. .35792	
	/rpt_family="MIR"	
	/rpt_family="L1"	
repeat_region	37865. .38147	
repeat_region	/rpt_family="Alu"	
repeat_region	complement(40345. .40621)	
repeat_region	/rpt_family="Alu"	
repeat_region	complement(40641. .40724)	
repeat_region	/rpt_family="L1"	
repeat_region	complement(41832. .42104)	
repeat_region	/rpt_family="L1"	
repeat_region	complement(42971. .43259)	
repeat_region	/rpt_family="Alu"	
repeat_region	43358. .43381	
	/note="(GT)12"	
	/rpt_type=tandem	
	/rpt_unit=GT	
repeat_region	44213. .44530	
repeat_region	/rpt_family="Alu"	
repeat_region	complement(45021. .45151)	
misc_feature	/rpt_family="Alu"	
misc_feature	45278. .45448	
repeat_region	/note="GRAIL 2 excellent exon, frame 2"	
repeat_region	47167. .47356	
repeat_region	/rpt_family="Alu"	
misc_feature	48025. .48095	
repeat_region	/note="GRAIL 2 excellent exon, frame 0"	
repeat_region	51462. .51636	
repeat_region	/rpt_family="MIR"	
prim_transcript	join(52201. .52421,56063. .56185)	
	/standard_names="v175d07"	
repeat_region	/note="84% identity mouse dbEST:AA498081"	
repeat_region	53603. .53735	
repeat_region	/rpt_family="MIR"	
repeat_region	53872. .54116	
repeat_region	/rpt_family="Alu"	
repeat_region	54133. .54168	
	/note="(CT)18"	
	/rpt_type=tandem	
	/rpt_unit=CT	
repeat_region	complement(55198. .55325)	
misc_feature	/rpt_family="MIR"	
misc_feature	56203. .56340	
prim_transcript	/note="GRAIL 2 excellent exon, frame 2"	
prim_transcript	join(56293. .56421,58623. .58786)	
misc_feature	/note="99% identity dbEST:AA095819"	
repeat_region	56941. .57036	
repeat_region	/note="GRAIL 2 excellent exon, frame 0"	
repeat_region	complement(57325. .57413)	
prim_transcript	/rpt_family="MIR"	
prim_transcript	join(58623. .58971,59089. .59471)	
	/standard_names="yq85c07"	
repeat_region	/note="98% identity dbEST:H53759"	
repeat_region	59788. .60079	
repeat_region	/rpt_family="Alu"	
repeat_region	60852. .60891	
	/note="(AAG)10"	
	/rpt_type=tandem	
repeat_region	/rpt_unit=AAAG	
repeat_region	60889. .60940	
	/note="(AAG)13"	
	/rpt_type=tandem	
repeat_region	/rpt_unit=AAAG	
repeat_region	60955. .60978	
	/note="(AAG)6"	
	/rpt_type=tandem	
repeat_region	/rpt_unit=AAAG	
repeat_region	60978. .61049	
	/note="(GAAA)18"	
	/rpt_type=tandem	
repeat_region	/rpt_unit=GAAA	
repeat_region	complement(62005. .62123)	
repeat_region	/rpt_family="Alu"	
repeat_region	62514. .62638	
	/rpt_family="MIR"	

```

repeat_region 65667..65690
/note="(GCT)8"
/rpt_type=tandem
/rpt_unit=GCT
complement(65752..66069)
/rpt_family="Alu"
repeat_region 68284..68370
/rpt_family="MIR"
repeat_region 68555..68858
/rpt_family="Alu"
repeat_region 69008..69149
/rpt_family="MIR"
misc_feature 69259..69355
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 70216..70236
/note="(T)21"
/rpt_type=tandem
/rpt_unit=T
repeat_region complement(70218..70419)
/rpt_family="Alu"
repeat_region 70953..70976
/note="(CT)12"
/rpt_type=tandem
/rpt_unit=CT
repeat_region 71078..71153
/rpt_family="MIR"
repeat_region complement(71422..71661)
/rpt_family="MIR"
repeat_region 71740..71986
/rpt_family="Alu"
repeat_region 71958..71986
/note="(A)29"
/rpt_type=tandem
/rpt_unit=A
repeat_region complement(73579..74130)

```

```

Query Match          79.5%; Score 32.6; DB 9; Length 112088;
Best Local Similarity 85.4%; Pred. NO. 0.13;
Matches 35; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGGA 41
    ||||| ||| |||||:||||| ||| |||||
Db 97204 CCTATCTCAACTAAATAACAAAAGCCAGGCATGGTGGA 97244

```

Search completed: February 11, 2003, 05:30:34  
 Job time : 277.088 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 21:06:50 ; Search time 141.088 Seconds  
(without alignments)  
8457.264 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_920\_960

Perfect score: 41

Sequence: 1 ctttgtgtggtgatttctt.....crtgtgtaatggtgcctcg 41

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_fod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	40.6	99.0	1680	6	AX394457	AX394457 Sequence
2	40.6	99.0	5503	9	HUMCYP2DG	M33189 Human debri
3	40.6	99.0	9432	6	AX394456	AX394456 Sequence
4	40.6	99.0	9432	9	HUMCYP2D6	M33388 Human cytoc
5	40.6	99.0	13278	9	HSCYP2D7A	X58467 Human CYP2D
6	40.6	99.0	13677	9	HSCYP2D7B	X58468 Human CYP2D
7	40.6	99.0	17060	9	HUMCYP8P	M33387 Human debri
8	40.6	99.0	114846	9	HS257120	AL021878 Human DNA
9	40.2	98.0	1669	6	AX207224	AX207224 Sequence
10	31.6	77.1	5884	6	AX345458	AX345458 Sequence
11	31.6	77.1	5884	6	AX348344	AX348344 Sequence
12	26.8	65.4	131664	2	AC097793	AC097793 Rattus no
13	26.8	65.4	188353	2	AC111843	AC111843 Rattus no
14	26.4	64.4	194184	2	AC111138	AC111138 Mus muscu
15	26.2	63.9	171777	2	AC073151	AC073151 Mus muscu
16	25.8	62.9	136016	2	AC108309	AC108309 Rattus no
17	25.8	62.9	164176	2	AC112075	AC112075 Rattus no
18	25.8	62.9	166654	2	AL671920	AL671920 Mus muscu
19	25.8	62.9	168738	2	AC121944	AC121944 Mus muscu
20	25.8	62.9	191931	2	AC120907	AC120907 Rattus no
21	25.8	62.9	204492	2	AC087560	AC087560 Mus muscu
22	25.8	62.9	226249	2	AC098624	AC098624 Rattus no
23	25.6	62.9	253913	2	AC126935	AC126935 Mus muscu
24	25.6	62.4	41724	3	CBRG27P14	AC084427 Caenorhab
25	25.2	61.5	1127	9	HS2QT100	296575 H.sapiens t
26	25.2	61.5	33318	2	AC131089	AC131089 Homo sapi
27	25.2	61.5	33901	2	AC127087	AC127087 Homo sapi
28	25.2	61.5	67512	2	AC117714	AC117714 Mus muscu
29	25.2	61.5	71219	2	AC101609	AC101609 Mus muscu
30	25.2	61.5	73557	2	AC093631	AC093631 Homo sapi
31	25.2	61.5	131277	10	AL669860	AL669860 Mouse DNA
32	25.2	61.5	131532	2	AC108967	AC108967 Rattus no
33	25.2	61.5	157979	2	AC106162	AC106162 Rattus no
34	25.2	61.5	164439	9	AL356794	AL356794 Human DNA
35	25.2	61.5	172811	2	AC103034	AC103034 Rattus no
36	25.2	61.5	175693	2	AC102241	AC102241 Mus muscu
37	25.2	61.5	178317	9	AC112081	AC112081 Homo sapi
38	25.2	61.5	182559	2	AC131097	AC131097 Homo sapi
39	25.2	61.5	182726	2	AL591426	AL591426 Mus muscu
40	25.2	61.5	183363	2	AL844183	AL844183 Mus muscu
41	25.2	61.5	195180	2	AL591146	AL591146 Mus muscu
42	25.2	61.5	195324	2	AC115912	AC115912 Mus muscu
43	25.2	61.5	204428	10	AL607074	AL607074 Mouse DNA
44	25.2	61.5	211804	2	AC094607	AC094607 Rattus no
45	25.2	61.5	240938	2	AL731665	AL731665 Mus muscu

ALIGNMENTS

RESULT 1  
AX394457  
LOCUS AX394457  
DEFINITION Sequence 2 from Patent WO0218638.  
ACCESSION AX394457  
VERSION AX394457.1 GI:21065595  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1  
REFERENCE  
AUTHORS Risinger, C., Andersson, M.K., Lewander, T. and Ollasson, E.  
TITLE Detection of cyp2d6 polymorphisms  
JOURNAL Patent: WO 0218638-A 2 07-MAR-2002;

linear PAT 18-MAY-2002

```

FEATURES             Gemini Genomics PLC (GB)
source               Location/Qualifiers
1..1680
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT          413 a 379 c 539 g 342 t 7 others
ORIGIN
1 CTTTGTGGGTGATTTCTGCGTGTGTAATCGTGCCCTG 41
|||||
Db 920 CTTTGTGGGTGATTTCTGCGTGTGTAATCGTGCCCTG 960

RESULT 2
HUMCYP2D6           5503 bp  DNA  linear  PRI 27-APR-1993
LOCUS               Human debrisoquine 4-hydroxylase mutant allele (CYP2D6-MA1) gene,
DEFINITION          complete cds.
ACCESSION           M33189.1 GI:181305
VERSION             debrisoquine 4-hydroxylase.
KEYWORDS            Human individual MAGA DNA.
SOURCE              Homo sapiens
ORGANISM            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS             Gonzalez,F.J.
JOURNAL             Unpublished (1990)
COMMENT             Draft entry and computer-readable sequence for [1] kindly submitted
                    by F.Gonzalez, 23-MAR-1990, for release after publication.
                    Author address: F.Gonzalez
                    National Cancer Institute
                    Bldg 37 Rm. 3E-24
                    National Institute of Health
                    Bethesda, Md 20892.
FEATURES             Location/Qualifiers
source              1..5503
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
TATA_signal         689..702
prim_transcript     726..5103
                    /note="debrisoquine 4-hydroxylase mRNA and introns"
CDS                join(814..993,1696..1877,2419..2571,2661..2820,3254..3430,
3621..3762,3970..4157,4612..4753,4852..5030)
                    /note="debrisoquine 4-hydroxylase"
                    /codon_start=1
                    /protein_id="AAA35737.1"
                    /db_xref="GI:181306"
                    /translation="MGLALVPLAVIVAIFLLVDMHRRQWAARYSPGPLPLPGLG
NLLHVDFTQYFCFLRRFGDFSLQLAWTPVVLNGLAAVREAMVTGCDTADRP
PVPTIQLFGPGRSOGKORGVEGALWARAEALLRLHLOLPGQOEAVAGVDRGG
RLPLCLRQLNRPFRPNGLDKAVSNVIALSLTCGRFEYDDPFLRLDLIAQGLKE
ESGFLREVLNAPVLLHLPALAGVLRFQKAFLTQDLELTHEHRMTWDPAPPRDLTE
AFLAMEKAKNGPSSFNENLRIVADLFSAGMVTSTTLAWGLLMLILHPDQRRV
QGEIDDVGGQVRPEMGDAQHPYTTAVIHEVQREGDIVPLGVTHMTSRDIEVGFR
PKGTLTINLSVLKDEAVKEKPFPHFHEFLDAQGHVKEAFLEPFSAGRACLGEP
LARLELFFFTSLQHFSEFSVPTQGRPSHHGVFAFLVTPSPYELCAVPR"
                    <814..993
                    /note="debrisoquine 4-hydroxylase"
                    /number=1
exon                994..1695
intron              /note="debrisoquine 4-hydroxylase intron A"
exon                1696..1877
                    /number=2
intron              1878..2418
exon                /note="debrisoquine 4-hydroxylase intron B"
                    2419..2571
                    /number=3
intron              2572..2660

```

```

/note="debrisoquine 4-hydroxylase intron C"
2661..2820
/number=4
intron              2821..3253
/note="debrisoquine 4-hydroxylase intron D"
3254..3430
/number=5
intron              3431..3620
/note="debrisoquine 4-hydroxylase intron E"
3621..3762
/number=6
intron              3763..3969
/note="debrisoquine 4-hydroxylase intron F"
3970..4157
/number=7
intron              4158..4611
/note="debrisoquine 4-hydroxylase intron G"
4612..4753
/number=8
intron              4754..4851
/note="debrisoquine 4-hydroxylase intron H"
4852..5030
/note="debrisoquine 4-hydroxylase"
/number=9
BASE COUNT          1066 a 1537 c 1851 g 1049 t
ORIGIN              Chromosome 22.
Query Match          99.0%; Score 40.6; DB 9; Length 5503;
Best Local Similarity 97.6%; Pred. No. 5.5e-06;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTGTGGGTGATTTCTGCGTGTGTAATCGTGCCCTG 41
|||||
Db 114 CTTTGTGGGTGATTTCTGCGTGTGTAATCGTGCCCTG 154

RESULT 3
AX394456            9432 bp  DNA  linear  PAT 18-MAY-2002
LOCUS               Sequence 1 from Patent WO0218638.
DEFINITION          AX394456
ACCESSION           AX394456
VERSION             AX394456.1 GI:21065594
KEYWORDS            human.
SOURCE              Homo sapiens
ORGANISM            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS             Risinger,C., Andersson,M.K., Lewander,T. and Ollasson,E.
JOURNAL             Detection of cyp2d6 polymorphisms
                    Patent: WO 0218638-A 1 07-MAR-2002;
                    Gemini Genomics PLC (GB)
FEATURES             Location/Qualifiers
source              1..9432
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
BASE COUNT          1964 a 2647 c 2976 g 1845 t
ORIGIN
Query Match          99.0%; Score 40.6; DB 6; Length 9432;
Best Local Similarity 97.6%; Pred. No. 5.3e-06;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CTTTGTGGGTGATTTCTGCGTGTGTAATCGTGCCCTG 41
|||||
Db 920 CTTTGTGGGTGATTTCTGCGTGTGTAATCGTGCCCTG 960

RESULT 4
HUMCYP2D6           9432 bp  DNA  linear  PRI 22-NOV-1994
LOCUS               Human cytochrome P450 IID6 (CYP2D6) gene, complete cds.
DEFINITION          HUMCYP2D6
ACCESSION           HUMCYP2D6

```

```
VERSION      M3388.1  GI:181303
KEYWORDS     cytochrome P450; cytochrome P450 IID6.
SOURCE       Human DNA, clone lambda2D-18/2.
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 9432)
AUTHORS      Kimura,S., Umeno,M., Skoda,R.C., Meyer,U.A. and Gonzalez,F.J.
TITLE        The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and
              identification of the polymorphic CYP2D6 gene, a related gene, and
              a pseudogene
JOURNAL      Am. J. Hum. Genet. 45 (6), 889-904 (1989)
MEDLINE      90072069
PUBMED       2574001
COMMENT      Draft entry and computer-readable sequence for [Am. J. Hum. Genet.
              45, 889-904 (1989)] kindly submitted
              by S.Kimura, 29-MAR-1990.
FEATURES     Location/Qualifiers
              1..9432
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /map="22q13.1"
              join(1532..1799,2503..2674,3225..3377,3466..3626,
              4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)
               /gene="CYP2D6"
              join(1532..1799,2503..2674,3225..3377,3466..3626,
              4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)
               /gene="CYP2D6"
               /product="cytochrome P450 IID6"
               /note="G00-132-127"
              1532..1799
               /partial
               /gene="CYP2D6"
               /note="cytochrome P450 IID6; G00-132-127"
               /number=1
              join(1620..1799,2503..2674,3225..3377,3466..3626,
              4060..4236,4427..4568,4776..4963,5418..5559,5658..5836)
               /gene="CYP2D6"
               /codon_start=1
               /product="cytochrome P450 IID6"
               /protein_id="AAAS3500.1"
               /db_xref="GI:181304"
               /db_xref="GDB:G00-132-127"
               /translation="MGELALPLAVIVAIFLLVLMHRRQWAARYPPGPLPLPLGLG
              NLLHVFQNTYCFDQLRRREGDFVSLQAWTPVVLNGLAAVREALVTHGEDTADRP
              PVPITQILGFGPSQGVFLARYGPWROQRFSVSTLRNLGLGKLSLEQWTEAAACL
              CAATANHSRPERENGLLDKAVSNVIAISLTGRRFEYDDPRFLRLDLAOGLEKESG
              FLREVLNAPVLLHIPALAGKVLRFQKAFLTQLDELLEHRTWDPQAPRDLTEAFL
              AEMEKAGNPESSENDELRIIVADLFSAGWVTSTTLAWGLLLMLHPVORRVOQE
              IDDIIGQVRRPEMGDQAHMPTTAVIHEVQRGDIVPLGVTHMTSDIEVQGFRIKPG
              TTLITNLSVLKDEAVWEKPFRIPEHFELDAQGHFKPEAFLPFSRAGCLGEPLAR
              MELFLFTSLQLHFSFSVPTGQPRSHHCVEAFILVSPSPYELCAVPR"
              1800..2502
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=1
              2503..2674
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=2
              2675..3224
               /gene="CYP2D6"
               /note="G00-132-127; does not fit consensus"
               /number=2
              3225..3377
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=3
              3378..3465
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=3
              3466..3626
               /number=4
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=4
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=5
              4237..4426
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=6
              4427..4568
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=7
              4776..4963
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=8
              5418..5559
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=9
              5560..5657
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=8
              5658..5909
               /partial
               /gene="CYP2D6"
               /note="G00-132-127"
               /number=9
              1964 a 2647 c 2976 g 1845 t
              BASE COUNT
              ORIGIN
              Query Match 99.0%; Score 40.6; DB 9; Length 9432;
              Best Local Similarity 97.6%; Pred. No. 5.3e-06;
              Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
              QY 1 CTTTGTGGTGGTGAATTTCTGCGTGTGTAATCGTGCCCTG 41
              |||||
              Db 920 CTTTGTGGTGGTGAATTTCTGCGTGTGTAATCGTGCCCTG 960
              |||||
              RESULT 5
              LOCUS      HSCYP2D7A 13278 bp DNA linear PRI 21-OCT-1992
              DEFINITION Human CYP2D7A pseudogene for cytochrome P450 2D6.
              ACCESSION X58467
              VERSION X58467.1 GI:30336
              KEYWORDS CYP2D7A gene; Cytochrome P450; cytochrome P450 2D6; pseudogene.
              SOURCE Homo sapiens.
              ORGANISM Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              REFERENCE 1 (bases 1 to 13278)
              AUTHORS Heim,M.H.
              TITLE Direct Submission
              JOURNAL Submitted (25-MAR-1991) M.H. Heim, Dept of Pharmacology, Biocentre
              University of Basel, Klingelbergstr 70, 4056 Basel, SWITZERLAND
              REFERENCE 2 (bases 1 to 13278)
```

AUTHORS Heim,M.H. and Meyer,U.A.  
TITLE Evolution of a highly polymorphic human cytochrome P450 gene  
JOURNAL Genomics 14 (1), 49-58 (1992)  
MEDLINE 93052308  
PUBMED 1358797  
COMMENT See X58468, and Am. J. Hum. Genet. 47:994-1001(1990).  
FEATURES  
source  
1. .13278  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clone="42"  
/clone\_lib="EMBL4"  
/dev\_stage="adult"  
1154..5489  
/gene="CYP2D7AP"  
Join(1154..1424,2125..2296,2823..2975,3064..3225,  
3651..3827,4020..4161,4356..4542,4998..5139,5238..5489)  
/gene="CYP2D7AP"  
1154..1424  
/gene="CYP2D7AP"  
/number=1  
Join(1242..1424,2125..2296,2823..2975,3064..3225,  
3651..3827,4020..4161,4356..4542,4998..5139,5238..5416)  
/gene="CYP2D7AP"  
/pseudo  
/codon\_start=1  
/db\_xref="PID:e34292"  
1425..2124  
/gene="CYP2D7AP"  
/number=1  
2125..2296  
/gene="CYP2D7AP"  
/number=2  
2297..2822  
/gene="CYP2D7AP"  
/number=2  
2823..2975  
/gene="CYP2D7AP"  
/number=3  
2976..3063  
/gene="CYP2D7AP"  
/number=3  
3064..3225  
/gene="CYP2D7AP"  
/number=4  
3226..3650  
/gene="CYP2D7AP"  
/number=4  
3651..3827  
/gene="CYP2D7AP"  
/number=5  
3828..4019  
/gene="CYP2D7AP"  
/number=5  
4020..4161  
/gene="CYP2D7AP"  
/number=6  
4162..4355  
/gene="CYP2D7AP"  
/number=6  
4356..4542  
/gene="CYP2D7AP"  
/number=7  
4543..4997  
/gene="CYP2D7AP"  
/number=7  
4998..5139  
/gene="CYP2D7AP"  
/number=8  
5140..5237  
/gene="CYP2D7AP"

exon  
/number=8  
5238..5489  
/gene="CYP2D7AP"  
/number=9  
7887..7908  
/note="ACCTTCCCC repeat"  
BASE COUNT 2902 a 3664 c 3968 g 2744 t  
ORIGIN  
Query Match 99.0%; Score 40.6; DB 9; Length 13278;  
Best Local Similarity 97.6%; Pred. No. 5.2e-06;  
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTGTGTGGTGATTCTTCTGRTGTGTAATCGTGTCCTG 41  
|||||  
Db 542 CTTTGTGTGGTGATTCTTCTGATGTGTAATCGTGTCCTG 582  
RESULT 6  
LOCUS HSCYP2D7B 13677 bp DNA linear PRI 21-OCT-1992  
DEFINITION Human CYP2D7BP pseudogene for cytochrome P450 2D6.  
ACCESSION X58468  
VERSION X58468.1 GI:30337  
KEYWORDS CYP2D7BP gene; Cytochrome P450; cytochrome P450 2D6; pseudogene.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Heim,M.H.  
Direct Submission  
Submitted (25-MAR-1991) M.H. Heim, Dept of Pharmacology, Biocentre  
University of Basel, Klingelbergstr 70, 4056 Basel, SWITZERLAND  
2 (bases 1 to 13677)  
Heim,M.H. and Meyer,U.A.  
Evolution of a highly polymorphic human cytochrome P450 gene  
cluster: CYP2D6  
Genomics 14 (1), 49-58 (1992)  
93052308  
PUBMED 1358797  
COMMENT See X58467, and Am. J. Hum. Genet. 47:994-1001(1990).  
FEATURES  
source  
1. .13677  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clone="45"  
/clone\_lib="EMBL4"  
/dev\_stage="adult"  
1534..5868  
/gene="CYP2D7BP"  
Join(1534..1801,2504..2675,3203..3355,3444..3605,  
4031..4207,4400..4540,4735..4922,5377..5518,5617..5868)  
/gene="CYP2D7BP"  
1534..1801  
/gene="CYP2D7BP"  
/number=1  
Join(1622..1801,2504..2675,3203..3355,3444..3605,  
4031..4207,4400..4540,4735..4922,5377..5518,5617..5795)  
/gene="CYP2D7BP"  
/pseudo  
/codon\_start=1  
/db\_xref="PID:e34293"  
1802..2503  
/gene="CYP2D7BP"  
/number=1  
2504..2675  
/gene="CYP2D7BP"  
/number=2  
2676..3202  
/gene="CYP2D7BP"  
/number=2  
intron  
exon  
intron

exon	3203..3355 /gene="CYP2D7BP" /number=3	3203..3355 /gene="CYP2D7BP" /number=3	COMMENT	Draft entry and computer-readable sequence for [1] kindly submitted by S.Kimura, 29-MAR-1990.
intron	3356..3443 /gene="CYP2D7BP" /number=3	3356..3443 /gene="CYP2D7BP" /number=3	FEATURES	Location/Qualifiers
exon	3444..3605 /gene="CYP2D7BP" /number=4	3444..3605 /gene="CYP2D7BP" /number=4	source	1..17060 /organism="Homo sapiens" /db_xref="taxon:9606"
intron	3606..4030 /gene="CYP2D7BP" /number=4	3606..4030 /gene="CYP2D7BP" /number=4	TATA_signal	1276..1282
exon	4031..4207 /gene="CYP2D7BP" /number=5	4031..4207 /gene="CYP2D7BP" /number=5	prim_transcript	1304..6570 /note="CYP2D8P mRNA and introns"
intron	4208..4399 /gene="CYP2D7BP" /number=5	4208..4399 /gene="CYP2D7BP" /number=5	gene	join(1392..1568,3189..3360,3907..4059,4148..4310,4758..4934,5121..5262,5467..5651,6101..6242,6339..6516) /gene="CYP2D8P"
exon	4301..4922 /gene="CYP2D7BP" /number=6	4301..4922 /gene="CYP2D7BP" /number=6	CDS	join(1392..1568,3189..3360,3907..4059,4148..4310,4758..4934,5121..5262,5467..5651,6101..6242,6339..6516) /gene="CYP2D8P"
intron	4923..5376 /gene="CYP2D7BP" /number=7	4923..5376 /gene="CYP2D7BP" /number=7	exon	/pseudocodon_start=1 /product="debrisoquine 4-hydroxylase"
exon	5377..5518 /gene="CYP2D7BP" /number=8	5377..5518 /gene="CYP2D7BP" /number=8	exon	<1392..1568 /gene="CYP2D8P"
intron	5519..5616 /gene="CYP2D7BP" /number=8	5519..5616 /gene="CYP2D7BP" /number=8	exon	/number=1
exon	5617..5868 /gene="CYP2D7BP" /number=9	5617..5868 /gene="CYP2D7BP" /number=9	exon	3189..3360 /gene="CYP2D8P"
repeat_region	8267..8306 /note="ACCCCTCCCC"	8267..8306 /note="ACCCCTCCCC"	exon	/number=2 /pseudocodon_start=1
BASE COUNT	3066 a 3775 c 4107 g 2729 t		exon	3907..4059 /gene="CYP2D8P"
ORIGIN			exon	/number=3 /pseudocodon_start=1
Query Match	99.0%; Score 40.6; DB 9; Length 13677;		exon	4148..4310 /gene="CYP2D8P"
Best Local Similarity	97.6%; Pred. No. 5.2e-06;		exon	/number=4 /pseudocodon_start=1
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			exon	4758..4934 /gene="CYP2D8P"
QY 1 CTTTGTGGGTGATTTCGTCGTGTGTAATCGTGTCCTG 41			exon	/number=5 /gene="CYP2D8P"
Db 922 CTTTGTGGGTGATTTCGTCGTGTGTAATCGTGTCCTG 962			exon	/pseudocodon_start=1
RESULT 7			exon	5121..5262 /gene="CYP2D8P"
HUMCYP8P			exon	/number=6 /pseudocodon_start=1
LOCUS	17060 bp DNA linear PRI 09-NOV-1994		exon	5467..5651 /gene="CYP2D8P"
DEFINITION	Human debrisoquine 4-hydroxylase (CYP2D8P) and (CYP2D7) pseudogenes complete sequences.		exon	/number=7 /pseudocodon_start=1
ACCESSION	M33387		exon	6101..6242 /gene="CYP2D8P"
VERSION	M33387.1 GI:181320		exon	/number=8 /pseudocodon_start=1
KEYWORDS	debrisoquine 4-hydroxylase.		exon	6339..>6516 /gene="CYP2D8P"
SOURCE	Human DNA, clones lambda-2D-A and lambda-2D-B.		exon	/number=9 /pseudocodon_start=1
ORGANISM	Homo sapiens		exon	11209..11215 /pseudocodon_start=1
REFERENCE	1 (bases 1 to 17060)		exon	11236..11571 /pseudocodon_start=1
AUTHORS	Kimura.S., Umeno.M., Skoda.R.C., Meyer,U.A. and Gonzalez,F.J.		exon	11571..11571 /pseudocodon_start=1
TITLE	The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene		exon	11571..11571 /pseudocodon_start=1
JOURNAL	Am J. Hum. Genet. 45 (6), 889-904 (1989)		exon	11571..11571 /pseudocodon_start=1
MEDLINE	90072069		exon	11571..11571 /pseudocodon_start=1
PUBMED	2574001		exon	11571..11571 /pseudocodon_start=1



misc\_feature /note="match: GSS: Em:B13982"  
25611..25853  
misc\_feature /note="match: GSS: Em:B14069"  
complement(26269..26446)  
misc\_feature /note="match: GSS: Em:AF046780"  
26890..26995  
repeat\_region /note="match: GSS: Em:A2083430"  
27147..28076  
repeat\_region /note="HERVI repeat: matches 4455..5408 of consensus"  
28386..30853  
repeat\_region /note="HERVI repeat: matches 2098..4455 of consensus"  
30843..30912  
repeat\_region /note="HERVI repeat: matches 2098..2163 of consensus"  
30902..30971  
repeat\_region /note="HERVI repeat: matches 2098..2163 of consensus"  
30961..31556  
repeat\_region /note="HERVI repeat: matches 1561..2163 of consensus"  
31873..33456  
repeat\_region /note="HERVI repeat: matches 9..1561 of consensus"  
33464..33959  
repeat\_region /note="L1P8a repeat: matches 1..510 of consensus"  
33960..34110  
repeat\_region /note="L1P8a repeat: matches 1226..1378 of consensus"  
34415..35367  
repeat\_region /note="L1P8a repeat: matches -656..1226 of consensus"  
37163..37323  
repeat\_region /note="Harlequin repeat: matches 912..1086 of consensus"  
37324..37437  
repeat\_region /note="57 copies 2 mer ga 79% conserved"  
complement(join(42379..42591,42916..43112))  
misc\_feature /note="match: STS: Em:G27508"  
complement(join(42491..42603,42916..42945))  
repeat\_region /note="match: STS: Em:G43129"  
43992..44915  
repeat\_region /note="PTR5 repeat: matches 40..93 of consensus"  
44916..45564  
repeat\_region /note="L1R12 repeat: matches 5..671 of consensus"  
48350..49467  
misc\_feature /note="CpG island"  
evidence=not\_experimental  
repeat\_region /note="17 copies 2 mer ca 100% conserved"  
54417..54450  
repeat\_region /note="MER77 repeat: matches 30..587 of consensus"  
54578..55083  
misc\_feature /note="CpG island"  
58051..59086  
evidence=not\_experimental  
misc\_feature /note="match: GSS: Em:AQ035975"  
65436..65550  
repeat\_region /note="23 copies 2 mer tt 82% conserved"  
65916..65961  
misc\_feature /note="match: STS: Em:G27630"  
66514..66549  
repeat\_region /note="18 copies 2 mer aa 80% conserved"  
66530..66802  
misc\_feature /note="match: GSS: Em:AQ617313"  
66577..66930  
misc\_feature /note="match: GSS: Em:AQ084812"  
66603..66893  
repeat\_region /note="match: GSS: Em:AQ044403"  
66974..67017  
misc\_feature /note="22 copies 2 mer cc 75% conserved"  
complement(69196..69374)  
repeat\_region /note="match: GSS: Em:B14383"  
69710..70103  
misc\_feature /note="match: GSS: Em:AQ181535"  
complement(78378..78961)  
misc\_feature /note="match: GSS: Em:AQ389013"  
complement(78421..78533)  
misc\_feature /note="match: GSS: Em:AQ042556"  
complement(78451..78977)  
misc\_feature /note="match: GSS: Em:AQ533248"

misc\_feature complement(81846..82274)  
/note="match: GSS: Em:AQ618257"  
repeat\_region 82587..83003  
/note="L1MA5A repeat: matches 5857..6292 of consensus"  
repeat\_region 83609..83702  
/note="MER21B repeat: matches 703..787 of consensus"  
repeat\_region 83697..83797  
/note="MER21B repeat: matches 686..794 of consensus"  
repeat\_region 83910..84357  
/note="MER21B repeat: matches 76..540 of consensus"  
repeat\_region 85375..85430  
/note="28 copies 2 mer aa 76% conserved"  
complement(90629..90900)  
misc\_feature /note="match: GSS: Em:B13983"  
repeat\_region 96343..96364  
misc\_feature /note="11 copies 2 mer ac 100% conserved"  
99274..99444  
repeat\_region /note="Other..Weak data"  
104028..104091  
/note="32 copies 2 mer aa 67% conserved"  
complement(107923..108378)  
misc\_feature /note="match: GSS: Em:AQ572846"  
Query Match 99.0%; Score 40.6; DB 9; Length 114846;  
Best Local Similarity 97.6%; Pred. No. 4.6e-06;  
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTGTGGTGATTTTCGTCRTGTGTAATCGTCCCTG 41  
|||||  
Db 51053 CTTTGTGGTGATTTTCGTCATGTGTAATCGTCCCTG 51013  
RESULT 9  
AX207224  
LOCUS AX207224 1669 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 1 from Patent WO0155432.  
ACCESSION AX207224  
VERSION AX207224.1 GI:15394976  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 1669)  
AUTHORS Raimundo,S. and Zanger,U.  
TITLE Polymorphisms in the human cyp2d6 gene promoter region and their  
use in diagnostic and therapeutic applications  
JOURNAL Patent: WO 0155432-A 1 02-AUG-2001;  
Epidaurus Biotechnologie AG (DE)  
FEATURES  
source  
1..1669  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="artificial sequence"  
BASE COUNT 413 a 376 c 534 g 338 t 8 others  
ORIGIN  
Query Match 98.0%; Score 40.2; DB 6; Length 1669;  
Best Local Similarity 95.1%; Pred. No. 8.5e-06;  
Matches 39; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CTTTGTGGTGATTTTCGTCRTGTGTAATCGTCCCTG 41  
|||||  
Db 920 CTTTGTGGTGATTTTCGTCATGTGTAATCGTCCCTG 960  
RESULT 10  
AX345458  
LOCUS AX345458 5884 bp DNA linear PAT 01-FEB-2002  
DEFINITION Sequence 529 from Patent WO0200928.  
ACCESSION AX345458  
VERSION AX345458.1 GI:18493344  
KEYWORDS  
SOURCE synthetic construct.

```

ORGANISM synthetic construct
REFERENCE artificial sequences.
1
AUTHORS Olek A., Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 529 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
source
1. 5884
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1259 a 92 c 1514 g 3019 t
ORIGIN
Query Match 77.1%; Score 31.6; DB 6; Length 5884;
Best Local Similarity 85.0%; Pred. No. 0.028;
Matches 34; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 2 TTTGTGGGTGATTTCTGTCGTGTAATCGTGTCCTCG 41
|||||
Db 4302 TTTGTGGGTGATTTCTGTCGTGTAATCGTGTCCTCG 4341
|||||

RESULT 11
AX348344 5884 bp DNA linear PAT 06-FEB-2002
LOCUS
DEFINITION Sequence 39 from Patent WO0202806.
ACCESSION AX348344
VERSION AX348344.1 GI:18614380
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
1
REFERENCE
AUTHORS Olek A., Piepenbrock, C. and Berlin, K.
TITLE Method and nucleic acids for pharmacogenomic methylation analysis
JOURNAL Patent: WO 0202806-A 39 10-JAN-2002;
Epigenomics AG (DE)
FEATURES
source
1. 5884
/organism="synthetic construct"
/db_xref="taxon:32630"
/notes="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1259 a 92 c 1514 g 3019 t
ORIGIN
Query Match 77.1%; Score 31.6; DB 6; Length 5884;
Best Local Similarity 85.0%; Pred. No. 0.028;
Matches 34; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 2 TTTGTGGGTGATTTCTGTCGTGTAATCGTGTCCTCG 41
|||||
Db 4302 TTTGTGGGTGATTTCTGTCGTGTAATCGTGTCCTCG 4341
|||||

RESULT 12
AC097793/c
LOCUS
DEFINITION Rattus norvegicus clone CH230-94L22, *** SEQUENCING IN PROGRESS
ACCESSION AC097793
VERSION AC097793.4 GI:21723633
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus
REFERENCE
1 (bases 1 to 131664)
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Alsbrooks, S.L., Amarantunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davilla, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Homi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenwo, S., Oguh, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 131664)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 131664)
Worley, K.C.
Direct Submission
Submitted (11-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 10, 2002 this sequence version replaced gi:20335376.
-----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: GFFU
Center clone name: CH230-94L22
-----
Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 90706 bases at least Q40
Consensus quality: 95513 bases at least Q30
Consensus quality: 99840 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 56 contigs. The true order of the pieces
* is not known and their order in this sequence record is

```



\* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 1287: contig of 1287 bp in length  
 \* 1288 1387: gap of unknown length  
 \* 1388 2360: contig of 1173 bp in length  
 \* 2561 2660: gap of unknown length  
 \* 2661 4119: contig of 1459 bp in length  
 \* 4120 4219: gap of unknown length  
 \* 4220 5266: contig of 1047 bp in length  
 \* 5267 5366: gap of unknown length  
 \* 5367 6596: contig of 1230 bp in length  
 \* 6597 6696: gap of unknown length  
 \* 6697 8127: contig of 1431 bp in length  
 \* 8128 8227: gap of unknown length  
 \* 8228 9420: contig of 1193 bp in length  
 \* 9421 9520: gap of unknown length  
 \* 9521 10631: contig of 1111 bp in length  
 \* 10632 10731: gap of unknown length  
 \* 10732 12226: contig of 1495 bp in length  
 \* 12227 12326: gap of unknown length  
 \* 12327 14139: contig of 1813 bp in length  
 \* 14140 14239: gap of unknown length  
 \* 14240 15930: contig of 1691 bp in length  
 \* 15931 16030: gap of unknown length  
 \* 16031 17060: contig of 1030 bp in length  
 \* 17061 17160: gap of unknown length  
 \* 17161 18443: contig of 1283 bp in length  
 \* 18444 18543: gap of unknown length  
 \* 18544 20121: contig of 1577 bp in length  
 \* 20122 20220: gap of unknown length  
 \* 20221 21309: contig of 1089 bp in length  
 \* 21310 21409: gap of unknown length  
 \* 21410 23479: contig of 2070 bp in length  
 \* 23480 23579: gap of unknown length  
 \* 23580 25370: contig of 1791 bp in length  
 \* 25371 25470: gap of unknown length  
 \* 25471 27568: contig of 2098 bp in length  
 \* 27569 27668: gap of unknown length  
 \* 27669 29458: contig of 1790 bp in length  
 \* 29459 29558: gap of unknown length  
 \* 29559 31357: contig of 1799 bp in length  
 \* 31358 31457: gap of unknown length  
 \* 31458 32469: contig of 1012 bp in length  
 \* 32470 32569: gap of unknown length  
 \* 32570 33858: contig of 1289 bp in length  
 \* 33859 33958: gap of unknown length  
 \* 33959 35636: contig of 1678 bp in length  
 \* 35637 35736: gap of unknown length  
 \* 35737 37109: contig of 1373 bp in length  
 \* 37110 37209: gap of unknown length  
 \* 37210 39170: contig of 1961 bp in length  
 \* 39171 39270: gap of unknown length  
 \* 39271 41002: contig of 1732 bp in length  
 \* 41003 41102: gap of unknown length  
 \* 41103 43242: contig of 2140 bp in length  
 \* 43243 43342: gap of unknown length  
 \* 43343 45259: contig of 1917 bp in length  
 \* 45260 45359: gap of unknown length  
 \* 45360 46688: contig of 1329 bp in length  
 \* 46689 46788: gap of unknown length  
 \* 46789 48442: contig of 1654 bp in length  
 \* 48443 51106: contig of 2564 bp in length  
 \* 51107 51206: gap of unknown length  
 \* 51207 52923: contig of 1717 bp in length  
 \* 52924 53023: gap of unknown length  
 \* 53024 54846: contig of 1823 bp in length  
 \* 54847 54947: gap of unknown length  
 \* 54948 56770: contig of 1824 bp in length  
 \* 56771 56870: gap of unknown length

\* 56871 58716: contig of 1846 bp in length  
 \* 58717 58816: gap of unknown length  
 \* 58817 60243: contig of 1427 bp in length  
 \* 60244 60343: gap of unknown length  
 \* 60344 62884: contig of 2541 bp in length  
 \* 62885 62984: gap of unknown length  
 \* 62985 65401: contig of 2417 bp in length  
 \* 65402 65501: gap of unknown length  
 \* 65502 68569: contig of 3068 bp in length  
 \* 68570 68669: gap of unknown length  
 \* 68670 69865: contig of 1196 bp in length  
 \* 69866 69965: gap of unknown length  
 \* 69966 72046: contig of 2081 bp in length  
 \* 72047 72146: gap of unknown length  
 \* 72147 74764: contig of 2618 bp in length  
 \* 74765 74864: gap of unknown length  
 \* 74865 77130: contig of 2266 bp in length  
 \* 77131 77230: gap of unknown length  
 \* 77231 79178: contig of 1948 bp in length  
 \* 79179 79278: gap of unknown length  
 \* 79279 81579: contig of 2301 bp in length  
 \* 81580 81679: gap of unknown length  
 \* 81680 83756: contig of 2077 bp in length  
 \* 83757 83856: gap of unknown length  
 \* 83857 86861: contig of 2805 bp in length  
 \* 86862 86961: gap of unknown length  
 \* 86962 90333: contig of 3572 bp in length  
 \* 90334 90433: gap of unknown length  
 \* 90434 93651: contig of 3218 bp in length  
 \* 93652 93751: gap of unknown length  
 \* 93752 97171: contig of 3420 bp in length  
 \* 97172 97271: gap of unknown length  
 \* 97272 102426: contig of 5155 bp in length  
 \* 102427 102526: gap of unknown length  
 \* 102527 106613: contig of 4087 bp in length  
 \* 106614 106713: gap of unknown length  
 \* 106714 111362: contig of 4649 bp in length  
 \* 111363 111462: gap of unknown length  
 \* 111463 115510: contig of 4048 bp in length

## Query Match

Best Local Similarity 65.4%; Score 26.8; DB 2; Length 131664;

Matches 31; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 TTGTGTGGGTGATTTCTGTCRTGTGAATGTCGTCCTG 41

||||| ||||| || ||||| ||||| ||||| ||||| |||||

Db 84549 TTGTATGGGTGTTGTCTGCATGTCGTCTGTGTACCTG 84510

## RESULT 13

AC111843

LOCUS

AC111843 Rattus norvegicus clone CH230-277M14, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\*, 109 unordered pieces.

DEFINITION

AC111843

VERSION

AC111843.2 GI:21743812

KEYWORDS

HTG; HTGS\_PHASE1.

SOURCE

Norway rat.

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

1 (bases 1 to 188353)

AUTHORS

Muzny, D. M., Adams, C., Adio-Oduola, B., Ali-osman, F. R., Allen, C.,

Alsbrooks, S. L., Amaratunga, H. C., Are, J. R., Ayale, M., Banks, T.,

Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D.,

Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N. P.,

Buhay, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C.,

Carroll, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D.,

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,

Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R.,

Davila, M. L., Davis, C., Davy-Carroll, L., Dederich, D. A.,

Delaney, K. R., Deigado, O., Denny, A. L., Ding, Y., Dinh, H. H.,

Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J.,  
 Earnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamill, C. K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hale, S., Hamill, C. K.,  
 Hernandez, O., Hodgson, A., Hogues, M., Hollaway, C., Hollins, B.,  
 Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L. E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L.,  
 Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
 Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
 Massey, E., Mawhiney, E., McLeod, M. P., Meador, M., Mei, G., Metzker, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, B., Nwokwenkwo, S., Oguh, M., Okwuonu, G.,  
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y.,  
 Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
 Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,  
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
 Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,  
 Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstock, G. and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 188353)  
 Worley, K.C.  
 Direct Submission  
 Submitted (19-FEB-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 188353)  
 Worley, K.C.  
 Direct Submission  
 Submitted (17-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 13, 2002 this sequence version replaced gi:18701687.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GOIS  
 Center clone name: CH230-277M14  
 ----- Summary Statistics  
 Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 87903 bases at least Q40  
 Consensus quality: 95337 bases at least Q30  
 Consensus quality: 101488 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 109 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 1021: contig of 1021 bp in length  
 \* 1022 1121: gap of unknown length  
 \*

1122 2184: contig of 1063 bp in length  
 2185 2284: gap of unknown length  
 2285 3325: contig of 1041 bp in length  
 3326 3425: gap of unknown length  
 3426 4438: contig of 1013 bp in length  
 4439 4538: gap of unknown length  
 4539 5547: contig of 1009 bp in length  
 5548 5647: gap of unknown length  
 5648 6807: contig of 1160 bp in length  
 6808 6907: gap of unknown length  
 6908 7974: contig of 1067 bp in length  
 7975 8074: gap of unknown length  
 8075 9131: contig of 1057 bp in length  
 9132 9231: gap of unknown length  
 9232 10367: contig of 1036 bp in length  
 10368 10677: gap of unknown length  
 10678 11394: contig of 1027 bp in length  
 11395 11494: gap of unknown length  
 11495 12498: contig of 1004 bp in length  
 12499 12598: gap of unknown length  
 12599 13517: contig of 1019 bp in length  
 13518 13717: gap of unknown length  
 13718 14777: contig of 1060 bp in length  
 14778 14877: gap of unknown length  
 14878 15892: contig of 1015 bp in length  
 15893 15992: gap of unknown length  
 15993 17318: contig of 1326 bp in length  
 17319 17418: gap of unknown length  
 17419 18577: contig of 1159 bp in length  
 18578 18677: gap of unknown length  
 18678 19764: contig of 1087 bp in length  
 19765 19864: gap of unknown length  
 19865 21187: contig of 1323 bp in length  
 21188 21287: gap of unknown length  
 21288 22887: contig of 1600 bp in length  
 22888 22987: gap of unknown length  
 22988 24047: contig of 1060 bp in length  
 24048 24147: gap of unknown length  
 24148 25641: contig of 1494 bp in length  
 25642 25741: gap of unknown length  
 25742 26931: contig of 1190 bp in length  
 26932 27031: gap of unknown length  
 27032 28338: contig of 1307 bp in length  
 28339 28438: gap of unknown length  
 28439 29441: contig of 1003 bp in length  
 29442 29541: gap of unknown length  
 29542 30745: contig of 1204 bp in length  
 30746 30845: gap of unknown length  
 30846 32144: contig of 1299 bp in length  
 32145 32244: gap of unknown length  
 32245 33577: contig of 1333 bp in length  
 33578 33677: gap of unknown length  
 33678 35491: contig of 1814 bp in length  
 35492 35591: gap of unknown length  
 35592 36882: contig of 1291 bp in length  
 36883 36982: gap of unknown length  
 36983 38118: contig of 1136 bp in length  
 38119 38218: gap of unknown length  
 38219 39526: contig of 1308 bp in length  
 39527 39626: gap of unknown length  
 39627 40758: contig of 1132 bp in length  
 40759 40858: gap of unknown length  
 40859 42360: contig of 1502 bp in length  
 42361 42460: gap of unknown length  
 42461 43523: contig of 1063 bp in length  
 43524 43623: gap of unknown length  
 43624 45204: contig of 1581 bp in length  
 45205 45304: gap of unknown length  
 45305 46353: contig of 1049 bp in length  
 46354 46453: gap of unknown length  
 46454 47587: contig of 1134 bp in length  
 47588 47687: gap of unknown length  
 47688 48805: contig of 1118 bp in length



\* 6439 6538: gap of 100 bp  
\* 6539 8886: contig of 2348 bp in length  
\* 8887 8986: gap of 100 bp  
\* 8987 12841: contig of 3855 bp in length  
\* 12842 12941: gap of 100 bp  
\* 12942 15693: contig of 2752 bp in length  
\* 15694 15793: gap of 100 bp  
\* 15794 18535: contig of 2742 bp in length  
\* 18536 18635: gap of 100 bp  
\* 18636 21710: contig of 3075 bp in length  
\* 21711 21810: gap of 100 bp  
\* 21811 24877: contig of 3067 bp in length  
\* 24878 24977: gap of 100 bp  
\* 24978 29713: contig of 4736 bp in length  
\* 29714 29813: gap of 100 bp  
\* 29814 33920: contig of 4107 bp in length  
\* 33921 34020: gap of 100 bp  
\* 34021 40161: contig of 6141 bp in length  
\* 40162 40261: gap of 100 bp  
\* 40262 47448: contig of 7187 bp in length  
\* 47449 47548: gap of 100 bp  
\* 47549 55863: contig of 8315 bp in length  
\* 55864 55963: gap of 100 bp  
\* 55964 67005: contig of 11042 bp in length  
\* 67006 67105: gap of 100 bp  
\* 67106 114741: contig of 47636 bp in length  
\* 114742 114841: gap of 100 bp  
\* 114842 137081: contig of 22240 bp in length  
\* 137082 137181: gap of 100 bp  
\* 137182 181365: contig of 44184 bp in length  
\* 181366 181465: gap of 100 bp  
\* 181466 194184: contig of 12719 bp in length.  
FEATURES  
source  
1. .194184  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP23-364J12"  
/clone\_lib="RPCI-23 Female Mouse BAC"  
1. .4349  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"  
4450. .6438  
/note="assembly\_fragment"  
6539. .8886  
/note="assembly\_fragment"  
8987. .12841  
/note="assembly\_fragment"  
12942. .15693  
/note="assembly\_fragment"  
15794. .18535  
/note="assembly\_fragment"  
18636. .21710  
/note="assembly\_fragment"  
21811. .24877  
/note="assembly\_fragment"  
24978. .29713  
/note="assembly\_fragment"  
29814. .33920  
/note="assembly\_fragment"  
34021. .40161  
/note="assembly\_fragment"  
40262. .47448  
/note="assembly\_fragment"  
47549. .55863  
/note="assembly\_fragment"  
55964. .67005  
/note="assembly\_fragment"  
67106. .114741  
/note="assembly\_fragment"  
114842. .137081  
/note="assembly\_fragment"  
137182. .181365

misc\_feature /note="assembly\_fragment"  
181466. .194184  
/note="assembly\_fragment  
clone\_end:T7  
vector\_side:right"  
BASE COUNT 60921 a 35509 c 35745 g 60304 t 1705 others  
ORIGIN  
Query Match 64.4%; Score 26.4; DB 2; Length 194184;  
Best Local Similarity 78.9%; Pred. No. 3.2;  
Matches 30; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 4 TGTGTTGGTGATTTCTGTCRTGTGTAATCGTGCCTCG 41  
||||| ||| ||||| :||||| ||| |||||  
Db 146782 TGTGTCGGTGTCTTCGTGTGTGTAATGATGACCTG 146819  
RESULT 15  
AC073151  
LOCUS 171777 bp DNA linear HTG 15-MAY-2002  
DEFINITION Mus musculus clone RP23-234E13 strain C57BL6/J, WORKING DRAFT  
SEQUENCE, 32 unordered pieces.  
AC073151  
HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
VERSION AC073151.8 GI:15143516  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 171777)  
AUTHORS Montgomery,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,  
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,  
Gordon,M., Goltz,J.S. and Kucherlapati,R.  
High Throughput Mouse Sequencing  
Unpublished  
2 (bases 1 to 171777)  
Montgomery,K.T., Grills,G., Han,J., Lee,E., Long,J., Pomerantz,R.,  
Ioshikhes,I.P., Shim,C., Decker,J., Thomas,E., Perera,A.,  
Gordon,M., Goltz,J.S. and Kucherlapati,R.  
Direct Submission  
Submitted (09-JUN-2000) Department of Molecular Genetics, Albert  
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,  
Bronx, NY 10461, USA  
On Aug 9, 2001 this sequence version replaced gi:15027676.  
-----Genome Center  
Center: Harvard Partners Genome Center  
Center Code: HPGC  
Web site: <http://www.hpcgg.org/Sequence/mouse.html>  
Contact: [hpgc@mendel.mgh.harvard.edu](mailto:hpgc@mendel.mgh.harvard.edu)  
-----Summary Statistics  
Center project name: ABY  
Sequencing vector: pUC18; L08752  
Chemistry: Dye-terminator Big Dye; 100%  
\*Consensus quality: 162972 at least Q20  
\*Consensus quality: 160757 at least Q30  
\*Consensus quality: 157344 at least Q40  
Estimated insert size: agarose-FP - N/A  
\*\*Estimated insert size: 171137 - sum-of-contigs  
Quality coverage: agarose-FP - N/A  
Quality coverage: 10.2 x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 70835: contig of 70835 bp in length  
\* 70836 70855: gap of unknown length  
\* 70856 101830: contig of 30975 bp in length  
\* 101831 101850: gap of unknown length

\* 101851 116928: contig of 15078 bp in length  
\* 116929 116948: gap of unknown length  
\* 116949 125555: contig of 8607 bp in length  
\* 125556 125575: gap of unknown length  
\* 125576 134576: contig of 9001 bp in length  
\* 134577 134596: gap of unknown length  
\* 134597 142742: contig of 8146 bp in length  
\* 142743 142762: gap of unknown length  
\* 142763 147456: contig of 4694 bp in length  
\* 147457 147476: gap of unknown length  
\* 147477 148318: contig of 842 bp in length  
\* 148319 148338: gap of unknown length  
\* 148339 149189: contig of 851 bp in length  
\* 149190 149210: gap of unknown length  
\* 149210 150191: contig of 982 bp in length  
\* 150191 150212: gap of unknown length  
\* 150212 151291: contig of 1080 bp in length  
\* 151291 151311: gap of unknown length  
\* 151311 152187: contig of 876 bp in length  
\* 152187 152208: gap of unknown length  
\* 152208 153450: contig of 1243 bp in length  
\* 153451 153471: gap of unknown length  
\* 153471 153589: contig of 119 bp in length  
\* 153590 153609: gap of unknown length  
\* 153610 155047: contig of 1438 bp in length  
\* 155048 155067: gap of unknown length  
\* 155068 155736: contig of 669 bp in length  
\* 155737 155756: gap of unknown length  
\* 155757 156778: contig of 1022 bp in length  
\* 156779 156798: gap of unknown length  
\* 156799 158277: contig of 1479 bp in length  
\* 158278 158297: gap of unknown length  
\* 158298 159432: contig of 1135 bp in length  
\* 159433 159452: gap of unknown length  
\* 159453 160449: contig of 997 bp in length  
\* 160450 160469: gap of unknown length  
\* 160470 161240: contig of 771 bp in length  
\* 161241 161260: gap of unknown length  
\* 161261 162614: contig of 1354 bp in length  
\* 162615 162634: gap of unknown length  
\* 162635 163410: contig of 776 bp in length  
\* 163411 163430: gap of unknown length  
\* 163431 164196: contig of 766 bp in length  
\* 164197 164217: gap of unknown length  
\* 164218 165483: contig of 1267 bp in length  
\* 165484 165503: gap of unknown length  
\* 165504 166376: contig of 873 bp in length  
\* 166377 166396: gap of unknown length  
\* 166397 167558: contig of 1162 bp in length  
\* 167559 167578: gap of unknown length  
\* 167579 168470: contig of 892 bp in length  
\* 168471 168490: gap of unknown length  
\* 168491 168612: contig of 122 bp in length  
\* 168613 168632: gap of unknown length  
\* 168633 169024: contig of 1192 bp in length  
\* 169025 169845: gap of unknown length  
\* 169846 171133: contig of 1289 bp in length  
\* 171134 171153: gap of unknown length  
\* 171154 171777: contig of 624 bp in length.

FEATURES

source

1..171777  
/organism="Mus musculus"  
/strain="C57BL6/J"  
/db\_xref="taxon:10090"  
/clone="RP23-234E13"  
/sex="male"  
1..70835  
/note="assembly\_name:Contig91  
clone\_end:77  
vector\_side:right"  
70856..101830  
/note="assembly\_name:Contig90"  
101851..116928

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

/note="assembly\_name:Contig89"  
116949..125555  
/note="assembly\_name:Contig88  
clone\_end:sp6  
vector\_side:right"  
125576..134576  
/note="assembly\_name:Contig87"  
134597..142742  
/note="assembly\_name:Contig86"  
142763..147456  
/note="assembly\_name:Contig85"  
147477..148318  
/note="assembly\_name:Contig84"  
148339..149189  
/note="assembly\_name:Contig83"  
149210..150191  
/note="assembly\_name:Contig82"  
150212..151291  
/note="assembly\_name:Contig81"  
151312..152187  
/note="assembly\_name:Contig80"  
152208..153450  
/note="assembly\_name:Contig79"  
153471..153589  
/note="assembly\_name:Contig78"  
153610..155047  
/note="assembly\_name:Contig77"  
155068..155736  
/note="assembly\_name:Contig76"  
155757..156778  
/note="assembly\_name:Contig75"  
156799..158277  
/note="assembly\_name:Contig74"  
158298..159432  
/note="assembly\_name:Contig73"  
159453..160449  
/note="assembly\_name:Contig72"  
160470..161240  
/note="assembly\_name:Contig71"  
161261..162614  
/note="assembly\_name:Contig70"  
162635..163410  
/note="assembly\_name:Contig69"  
163431..164196  
/note="assembly\_name:Contig68"  
164217..165483  
/note="assembly\_name:Contig67"  
165504..166376  
/note="assembly\_name:Contig66"  
166397..167558  
/note="assembly\_name:Contig65"  
167579..168470  
/note="assembly\_name:Contig64"  
168491..168612  
/note="assembly\_name:Contig63"  
168633..169824  
/note="assembly\_name:Contig62"  
169845..171133  
/note="assembly\_name:Contig61"  
171154..171777  
/note="assembly\_name:Contig60"  
BASE COUNT 47301 a 38940 c 39469 g 45409 t 658 others  
ORIGIN

Query Match 63.9%; Score 26.2; DB 2; Length 171777;  
Best Local Similarity 75.6%; Pred.No.3.9;  
Matches 31; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 CTTTGTGTGGTGATTTCCTGCTGCTAATCGTGCCTG 41  
|||||

Db 138859 CTTTGTGTGTGATTTCCTGCTGCTGCTGCTG 138899  
|||||

Search completed: February 11, 2003, 05:35:01  
Job time : 237.088 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 21:06:50 ; Search time 141.088 Seconds  
(without alignments)  
8457.264 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_600\_640

Perfect score: 41  
Sequence: 1 agaaagcagtggaggagac.....acccctcagcgagcccgaggag 41

Scoring table: IDENTITY-NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pi:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pi:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rtd:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	40.6	99.0	1669	6	AX207224 Sequence
2	40.6	99.0	1680	6	AX394457 Sequence
3	40.6	99.0	9432	6	AX394456 Sequence
4	40.6	99.0	9432	9	M33388 Human cytoC
5	40.6	99.0	13677	9	H58468 Human CYP2D
6	34.2	83.4	114846	9	HS257120 Human DNA
7	26.2	63.9	163856	2	AC102388 Mus muscu
8	24.8	60.5	232603	2	AC127560 Mus muscu
9	24.6	60.0	612	9	BC012882 Homo sapi
10	24.6	60.0	2555	6	AX078286 Sequence
11	24.6	60.0	2636	6	AX156560 Sequence
12	24.6	60.0	3071	9	AL136564 Homo sapi
13	24.6	60.0	3188	9	AK056443 Homo sapi
14	24.6	60.0	3194	9	AJ305226 Homo sapi
15	24.6	60.0	3222	9	BC017191 Homo sapi
16	24.6	60.0	3299	9	BC017191 Homo sapi
17	24.6	60.0	5875	9	HS305227 Homo sapi
18	24.6	60.0	75429	2	AK074091 Homo sapi
19	24.6	60.0	89948	9	HS756G23 Mus muscu
20	24.6	60.0	185174	2	AC118184 Human DNA
21	24.6	60.0	189291	2	AC118184 Rattus no
22	24.6	60.0	190654	2	AC102326 Mus muscu
23	24.4	59.5	91261	9	AC127113 Rattus no
24	24.2	59.0	230050	1	AL627277 Salmonell
25	24	58.5	103259	2	AC012430 Homo sapi
26	24	58.5	266915	2	AC103312 Homo sapi
27	23.8	58.0	110000	2	LMFLCHR32_03
28	23.8	58.0	204875	2	AC124466 Mus muscu
29	23.6	57.6	127400	5	AF170340 Xenopus l
30	23.6	57.6	221372	2	AC131121 Mus muscu
31	23.6	57.1	77448	2	AC124474 Mus muscu
32	23.4	57.1	187101	9	AC069565 Homo sapi
33	23.4	56.6	59181	2	AL589862 Human DNA
34	23.2	56.6	59181	2	AC131234 Homo sapi
35	23.2	56.6	80350	2	AC106370 Rattus no
36	23.2	56.6	108634	2	AC020813 Mus muscu
37	23.2	56.6	110000	2	LMFLCHR32_23
38	23.2	56.6	129120	9	HS187B23 Continuation (24 o
39	23.2	56.6	132870	2	AL031280 Human DNA
40	23.2	56.6	153845	9	AC119629 Rattus no
41	23.2	56.6	170916	2	AC025426 Homo sapi
42	23.2	56.6	173713	2	AC127576 Mus muscu
43	23.2	56.6	176265	2	AC095840 Rattus no
44	23.2	56.6	179192	2	AC126634 Rattus no
45	23.2	56.6	180533	2	AC129104 Homo sapi
					AC135807 Homo sapi

ALIGNMENTS

RESULT 1	AX207224	AX207224	Sequence	1669 bp	DNA	linear	PAT 30-AUG-2001
AX207224	LOCUS	AX207224	Sequence	1 from Patent WO0155432.			
DEFINITION	AX207224	AX207224	Accession				
VERSION	AX207224.1	GI:15394976					
KEYWORDS							
SOURCE			synthetic construct.				
ORGANISM			synthetic construct				
REFERENCE			artificial sequences.				
AUTHORS			1 (bases 1 to 1669)				
TITLE			Raimundo, S. and Zanger, U.				
JOURNAL			Polymorphisms in the human cyp2d6 gene promoter region and their use in diagnostic and therapeutic applications				
			Patent: WO 0155432-A 1 02-AUG-2001;				

Epidauros Biotechnologie AG (DE)

FEATURES  
source

1. .1669  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="artificial sequence"

BASE COUNT 413 a 376 c 534 g 338 t 8 others

ORIGIN

Query Match 99.0%; Score 40.6; DB 6; Length 1669;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGAAGCAGTGGAGGAGGACRACCTCAGGCGCGGAG 41  
|||||  
Db 600 AGAAGCAGTGGAGGAGGACRACCTCAGGCGCGGAG 640

RESULT 2  
AX394457  
LOCUS AX394457 1680 bp DNA linear PAT 18-MAY-2002  
DEFINITION Sequence 2 from Patent WO0218638.  
ACCESSION AX394457  
VERSION AX394457.1 GI:21065595  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM

REFERENCE  
AUTHORS Risinger,C., Andersson,M.K., Lewander,T. and Ollasson,E.  
TITLE Detection of cyp2d6 polymorphisms  
JOURNAL Patent: WO 0218638-A 2 07-MAR-2002;  
Gemini Genomics PLC (GB)

FEATURES  
source

1. .1680  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 413 a 379 c 539 g 342 t 7 others  
ORIGIN

Query Match 99.0%; Score 40.6; DB 6; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGAAGCAGTGGAGGAGGACRACCTCAGGCGCGGAG 41  
|||||  
Db 600 AGAAGCAGTGGAGGAGGACRACCTCAGGCGCGGAG 640

RESULT 3  
AX394456  
LOCUS AX394456 9432 bp DNA linear PAT 18-MAY-2002  
DEFINITION Sequence 1 from Patent WO0218638.  
ACCESSION AX394456  
VERSION AX394456.1 GI:21065594  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM

REFERENCE  
AUTHORS Risinger,C., Andersson,M.K., Lewander,T. and Ollasson,E.  
TITLE Detection of cyp2d6 polymorphisms  
JOURNAL Patent: WO 0218638-A 1 07-MAR-2002;  
Gemini Genomics PLC (GB)

FEATURES  
source

1. .9432  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 1964 a 2647 c 2976 g 1845 t  
ORIGIN

Query Match

Best Local Similarity 97.6%; Pred. No. 0.00076;  
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGCAGTGGAGGAGGACRACCTCAGGCGCGGAG 41  
|||||  
Db 600 AGAAGCAGTGGAGGAGGACCTCAGGCGCGGAG 640

RESULT 4  
HUMCYP2D6  
LOCUS

DEFINITION Human cytochrome P450 IID6 (CYP2D6) gene, complete cds.  
ACCESSION M33388  
VERSION M33388.1 GI:181303  
KEYWORDS cytochrome P450; cytochrome P450 IID6.  
SOURCE Human DNA, clone lambda2D-18/2.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Kimura,S., Umeno,M., Skoda,R.C., Meyer,U.A. and Gonzalez,F.J.  
TITLE The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene  
JOURNAL Am. J. Hum. Genet. 45 (6), 889-904 (1989)  
MEDLINE 90072069  
PUBMED 2574001

COMMENT Draft entry and computer-readable sequence for [Am. J. Hum. Genet. 45, 889-904 (1989)] kindly submitted by S.Kimura, 29-MAR-1990.  
FEATURES  
Location/Qualifiers  
1. .9432  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="22q13.1"  
join(1532..1799,2503..2674,3225..3377,3466..3626,4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)  
/gene="CYP2D6"  
join(1532..1799,2503..2674,3225..3377,3466..3626,4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)  
/product="cytochrome P450 IID6"  
/note="G00-132-127"  
1532..1799  
/partial  
/gene="CYP2D6"  
/note="cytochrome P450 IID6; G00-132-127"  
/number=1  
join(1620..1799,2503..2674,3225..3377,3466..3626,4060..4236,4427..4568,4776..4963,5418..5559,5658..5836)  
/gene="CYP2D6"  
/codon\_start=1  
/product="cytochrome P450 IID6"  
/protein\_id="AAAS3500.1"  
/db\_xref="GI:181304"  
/db\_xref="GDB:G00-132-127"  
/translation="MGLEALVPLAVIIVAFLLVLDLMHRRQRWAARYPPGGLPLPLGLG  
NLLHVDQNTIPYCFDQURRFQGVFSQLAWTPVVLNGLAARVREALVTHGEDTADRP  
PVPITQLIGFPGRSQGVFLARYGPWREORFVSSTLRNLGLGKKSLEQWTEAAACL  
CAAFANSHGSRFFRPNGLDRAVSNVIASLTCGRFRFYDDPRFLRLDLAEGLEESG  
FLREVLNAPVLLHIFALAGKVLRFQKAFLTQDLELTHRTWDPAPPRDLTEAFL  
AEMEKAKNPESSENDELRIIVADLFSGAMVTTSTTLANGLLMLLHPDVRVQGE  
IDDVIGVRRPENGDOAHMPYTTAVIHEVOREGDIPLGVTHTSRIEIVQGRIPKQ  
TTLITNLSVLKDEAVWEKFRFHPHFLOAQGHVFAPEALPFSAPRACGLPELAR  
MELFLFTLSLQHFVFSVPTGQPRRSHHGVAFLVSPSYELCAVPR"  
1800..2502  
/gene="CYP2D6"  
/note="G00-132-127"  
/number=1  
2503..2674  
/gene="CYP2D6"  
/note="G00-132-127"



intron  
/number-2  
2675..3224  
/gene="CYP2D6"  
/note="G00-132-127; does not fit consensus"  
/number-2  
3225..3377  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-3  
3378..3465  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-3  
3466..3626  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-4  
3627..4059  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-4  
4060..4236  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-5  
4237..4426  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-5  
4427..4568  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-6  
4569..4775  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-6  
4776..4963  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-7  
4964..5417  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-7  
5418..5559  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-8  
5560..5657  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-8  
5658..5909  
/partial  
/gene="CYP2D6"  
/note="G00-132-127"  
/number-9

BASE COUNT 1964 a 2647 c 2976 g 1845 t  
ORIGIN

Query Match 99.0%; Score 40.6; DB 9; Length 9432;  
Best Local Similarity 97.6%; Pred. No. 0.00076;  
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGCACTGGAGGAGGACACCCCTCAGGCAGCCCGGAG 41  
|||||

DB 600 AGAAGCACTGGAGGAGGACACCCCTCAGGCAGCCCGGAG 640  
|||||

RESULT 5  
HSCYP2D7B  
LOCUS

HSCYP2D7B 13677 bp DNA linear PRI 21-OCT-1992

DEFINITION Human CYP2D7BP pseudogene for cytochrome P450 2D6.  
ACCESSION X58468  
VERSION X58468.1 GI:30337  
KEYWORDS CYP2D7BP gene; Cytochrome P450; cytochrome P450 2D6; pseudogene.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 13677)  
AUTHORS Heim,M.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-MAR-1991) M.H. Heim, Dept of Pharmacology, Biocentre  
University of Basel, Klingelbergstr 70, 4056 Basel, SWITZERLAND  
REFERENCE 2 (bases 1 to 13677)  
AUTHORS Heim,M.H. and Meyer,U.A.  
TITLE Evolution of a highly polymorphic human cytochrome P450 gene  
JOURNAL Genomics 14 (1), 49-58 (1992)  
MEDLINE 93052308  
PUBMED 1358797  
COMMENT See X58467, and Am. J. Hum. Genet. 47:994-1001(1990).  
FEATURES  
source  
1..13677  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clone="45"  
/clone.lib="EMBL4"  
/dev\_stage="adult"  
1534..5868  
/gene="CYP2D7BP"  
Join(1534..1801,2504..2675,3203..3355,3444..3605,  
4031..4207,4400..4540,4735..4922,5377..5518,5617..5868)  
/gene="CYP2D7BP"  
1534..1801  
/gene="CYP2D7BP"  
/number-1  
Join(1622..1801,2504..2675,3203..3355,3444..3605,  
4031..4207,4400..4540,4735..4922,5377..5518,5617..5795)  
/gene="CYP2D7BP"  
/pseudo  
/codon\_start=1  
/db\_xref="PID:e34293"  
1802..2503  
/gene="CYP2D7BP"  
/number-1  
2504..2675  
/gene="CYP2D7BP"  
/number-2  
2676..3202  
/gene="CYP2D7BP"  
/number-2  
3203..3355  
/gene="CYP2D7BP"  
/number-3  
3356..3443  
/gene="CYP2D7BP"  
/number-3  
3444..3605  
/gene="CYP2D7BP"  
/number-4  
3606..4030  
/gene="CYP2D7BP"  
/number-4  
4031..4207  
/gene="CYP2D7BP"  
/number-5  
4208..4399  
/gene="CYP2D7BP"  
/number-5  
4400..4540  
/gene="CYP2D7BP"  
/number-6  
gene  
mRNA  
exon  
CDS  
intron  
exon  
intron  
exon  
intron  
exon  
intron  
exon  
intron  
exon  
intron  
exon

```

intron      4541..4734
            /gene="CYP2D7BP"
            /number=6
exon        4735..4922
            /gene="CYP2D7BP"
            /number=7
intron      4923..5376
            /gene="CYP2D7BP"
            /number=7
exon        5377..5518
            /gene="CYP2D7BP"
            /number=8
intron      5519..5616
            /gene="CYP2D7BP"
            /number=8
exon        5617..5868
            /gene="CYP2D7BP"
            /number=9
repeat_region 8267..8306
            /note="ACCTTCCCC"

BASE COUNT 3066 a 3775 c 4107 g 2729 t
ORIGIN

Query Match          99.0%; Score 40.6; DB 9; Length 13677;
Best Local Similarity 97.6%; Pred. No. 0.00071;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGAAGCAGTGGAGGAGGACACCTTCAGCGACCCCGGGAG 41
   |||||
Db 602 AGAAGCAGTGGAGGAGGACACCTTCAGCGACCCCGGGAG 642

RESULT 6
HS257120/c
LOCUS      HS257120          114846 bp    DNA    linear    PRI 22-NOV-2001
DEFINITION Human DNA sequence from clone RP1-257120 on chromosome
22q13.1-13.2, complete sequence.
ACCESSION AL021878
VERSION   AL021878.2  GI:17065905
KEYWORDS  HTG.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           Bridgeman,A.
           Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
           Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
           humquery@sanger.ac.uk
           On Nov 25, 2001 this sequence version replaced gi:3204432.
           During sequence assembly data is compared from overlapping clones.
           Where differences are found these are annotated as variations
           together with a note of the overlapping clone name. Note that the
           variation annotation may not be found in the sequence submission
           corresponding to the overlapping clone, as we submit sequences with
           only a small overlap as described above.
           The following abbreviations are used to associate primary accession
           numbers given in the feature table with their source databases:
           Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information
           on the WORMPEP database can be found at
           http://www.sanger.ac.uk/projects/C.elegans/wormpep
           This sequence
           was finished as follows unless otherwise noted: all regions were
           either double-stranded or sequenced with an alternate chemistry or
           covered by high quality data (i.e., phred quality >= 30); an
           attempt was made to resolve all sequencing problems, such as
           compressions and repeats; all regions were covered by at least one
           plasmid subclone or more than one M13 subclone; and the assembled
           was confirmed by restriction digest. This sequence was generated
           from part of bacterial clone contigs of human chromosome 22,
           constructed by the Sanger Centre Chromosome 22 Mapping Group.
           Further information can be found at
           http://www.sanger.ac.uk/HGP/Chr22

```

RP1-257120 is from the library RPCI-1 constructed by the group of  
 pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pCYPAC2  
 This sequence is the entire insert of clone RP1-257120 The true  
 right end of clone RP1-18601 is at 20171 in this sequence.

FEATURES  
 Location/Qualifiers  
 1..114846  
 /organism="Homo sapiens"  
 /db\_xref="RZPD:RPCI704120257"  
 /db\_xref="taxon:9606"  
 /chromosomes="22"  
 /map="q13.1-13.2"  
 /clone="RP1-257120"  
 /clone\_lib="RPCI-1"  
 1093..1682  
 /note="match: GSS: Em:AQ887153"  
 1097..1600  
 /note="match: GSS: Em:AQ563517"  
 complement(3894..4138)  
 /note="match: GSS: Em:AQ140240"  
 4148..4647  
 /note="match: GSS: Em:AQ684484"  
 7162..7366  
 /note="match: STS: Em:HS324WC5"  
 7357..7396  
 /note="20 copies 2 mer ac 82% conserved"  
 7369..7551  
 /note="match: STS: Em:HS324WC5"  
 7972..9283  
 /note="CpG island"  
 /evidence=not\_experimental  
 9981..10378  
 /note="match: GSS: Em:AQ124532"  
 13208..14275  
 /note="MER11C repeat: matches 1..1057 of consensus"  
 16086..16507  
 /note="L1MD repeat: matches 1552..1964 of consensus"  
 17055..17127  
 /note="L1MD repeat: matches 1482..1552 of consensus"  
 17434..17688  
 /note="L1MD repeat: matches 1222..1482 of consensus"  
 19152..19153  
 /note="clone RP1-18601  
 aa in this entry  
 deletion"  
 /replace="aaaaa"  
 19896..19939  
 /note="L1P repeat: matches 2416..2459 of consensus"  
 20246..20284  
 /note="L1P repeat: matches 2378..2416 of consensus"  
 20594..20660  
 /note="L1P repeat: matches 2311..2378 of consensus"  
 20965..20982  
 /note="L1P repeat: matches 2294..2311 of consensus"  
 21270..22130  
 /note="L1P repeat: matches 1445..2294 of consensus"  
 22425..22499  
 /note="L1P repeat: matches 1370..1445 of consensus"  
 22502..22994  
 /note="L1R10B repeat: matches 1..510 of consensus"  
 23020..25379  
 /note="BacV-int repeat: matches 4243..6625 of consensus"  
 25398..25662  
 /note="match: STS: Em:G07321"  
 25597..26083  
 /note="match: GSS: Em:BL3982"  
 25611..25853  
 /note="match: GSS: Em:BL4069"  
 complement(26269..26446)  
 /note="match: GSS: Em:AF046780"  
 26890..26995  
 /note="match: GSS: Em:A2083430"



Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zemek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 21, 2002 this sequence version replaced gi:17061474.  
All repeats were identified using RepeatMasker:  
Smith, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR  
Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information

Center project name: L18672  
Center clone name: 298.F.19

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 157585 bases at least Q40  
Consensus quality: 160552 bases at least Q30  
Consensus quality: 161468 bases at least Q20  
Insert size: 164000; agarose-1p  
Insert size: 162056; sum-of-contigs  
Quality coverage: 6.7 in Q20 bases; agarose-1p  
Quality coverage: 6.8 in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 19 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1  
\* 211 310: contig of 210 bp in length  
\* 311 gap of 100 bp  
\* 1491 1491: contig of 1181 bp in length  
\* 1492 1591: gap of 100 bp  
\* 1592 2716: contig of 1125 bp in length  
\* 2717 2816: gap of 100 bp  
\* 2817 3903: contig of 1087 bp in length  
\* 3904 4003: gap of 100 bp  
\* 4004 6451: contig of 2448 bp in length  
\* 6452 6551: gap of 100 bp  
\* 6552 8978: contig of 2427 bp in length  
\* 8979 9078: gap of 100 bp  
\* 9079 11520: contig of 2442 bp in length  
\* 11521 11620: gap of 100 bp  
\* 11621 13781: contig of 2161 bp in length  
\* 13782 13881: gap of 100 bp  
\* 13882 16379: contig of 2498 bp in length  
\* 16380 16479: gap of 100 bp  
\* 16480 19552: contig of 3073 bp in length  
\* 19553 19652: gap of 100 bp  
\* 19653 24331: contig of 4679 bp in length  
\* 24332 24431: gap of 100 bp  
\* 24432 29711: contig of 5280 bp in length  
\* 29712 29811: gap of 100 bp  
\* 35804: contig of 5993 bp in length

\* 35805 35904: gap of 100 bp  
\* 35905 43468: contig of 7564 bp in length  
\* 43469 43568: gap of 100 bp  
\* 43569 75257: contig of 31689 bp in length  
\* 75258 75357: gap of 100 bp  
\* 75358 101457: contig of 26100 bp in length  
\* 101458 101557: gap of 100 bp  
\* 101558 126283: contig of 24726 bp in length  
\* 126284 126383: gap of 100 bp  
\* 126384 155823: contig of 29440 bp in length  
\* 155824 155923: gap of 100 bp  
\* 155924 163856: contig of 7933 bp in length.

FEATURES  
Location/Qualifiers  
1. .163856

/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP24-298F19"  
/clone\_lib="RPC1-24 Male Mouse BAC"  
1. .210  
/note="assembly\_fragment"

misc\_feature

311. .1491  
/note="assembly\_fragment"

misc\_feature

1592. .2716  
/note="assembly\_fragment"

misc\_feature

2817. .3903  
/note="assembly\_fragment"

misc\_feature

4004. .6451  
/note="assembly\_fragment"

misc\_feature

6552. .8978  
/note="assembly\_fragment"

misc\_feature

9079. .11520  
/note="assembly\_fragment"

misc\_feature

11621. .13781  
/note="assembly\_fragment"

misc\_feature

13882. .16379  
/note="assembly\_fragment"

misc\_feature

16480. .19552  
/note="assembly\_fragment"

misc\_feature

19653. .24331  
/note="assembly\_fragment"

misc\_feature

24432. .29711  
/note="assembly\_fragment"

misc\_feature

29812. .35804  
/note="assembly\_fragment"

misc\_feature

35905. .43468  
/note="assembly\_fragment"

misc\_feature

43569. .75257  
/note="assembly\_fragment"

misc\_feature

75358. .101457  
/note="assembly\_fragment"

misc\_feature

101558. .126283  
/note="assembly\_fragment"

misc\_feature

126384. .155823  
/note="assembly\_fragment"

misc\_feature

155924. .163856  
/note="assembly\_fragment"

misc\_feature

clone\_end:t7  
vector\_side:right"

BASE COUNT 41088 a 39986 c 39793 g 41185 t 1804 others  
ORIGIN

Query Match 63.9%; Score 26.2; DB 2; Length 163856;  
Best Local Similarity 79.5%; Pred. No. 34;  
Matches 31; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AGAAGCAGCTGAGGAGGACRACCTCAGGCAGCCGCGG 39  
||||| ||||| ||| || ||||| ||||| |||  
Db 123076 AGATAGAGTGAGCAGCGCCTACAGTCAGGCAGCCAGG 123114

RESULT 8  
AC127560  
LOCUS

AC127560 232603 bp DNA linear HTG 18-JUL-2002

DEFINITION	Mus musculus chromosome UNK clone RP24-267118, WORKING DRAFT
ACCESSION	SEQUENCE, 6 unordered pieces.
VERSION	AC127560
KEYWORDS	AC127560.2 GI:21903649
SOURCE	HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
ORGANISM	house mouse.
REFERENCE	Mus musculus
AUTHORS	Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL	McPherson,J.D. and Waterston,R.H.
AUTHORS	1 (bases 1 to 232603)
TITLE	The sequence of Mus musculus clone
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 232603)
AUTHORS	McPherson,J.D. and Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (17-JUL-2002) Genome Sequencing Center, 4444 Forest Park
REFERENCE	3 (bases 1 to 232603)
AUTHORS	McPherson,J.D. and Waterston,R.H.
TITLE	Direct Submission
JOURNAL	Submitted (18-JUL-2002) Genome Sequencing Center, 4444 Forest Park
COMMENT	On Jul 18, 2002 this sequence version replaced gi:21886990.
	----- Genome Center -----
	Center: Washington University Genome Sequencing Center
	Center code: WUGSC
	Web site: http://genome.wustl.edu/gsc/index.shtml
	Contact: submissions@watson.wustl.edu
	----- Project Information -----
	Center project name: M_BB0267118
	----- Summary Statistics -----
	Sequencing vector: M13; 0%
	Chemistry: Dye-primer ET; 0% of reads
	Assembly: Dye-terminator Big Dye; 100% of reads
	Consensus quality: 229492 bases at least Q40
	Consensus quality: 229842 bases at least Q30
	Consensus quality: 230029 bases at least Q20
	* NOTE: This is a 'working draft' sequence. It currently
	* consists of 6 contigs. The true order of the pieces
	* is not known and their order in this sequence record is
	* arbitrary. Gaps between the contigs are represented as
	* runs of N, but the exact sizes of the gaps are unknown.
	* This record will be updated with the finished sequence
	* as soon as it is available and the accession number will
	* be preserved.
	* 1 12398: contig of 12398 bp in length
	* 12399 12498: gap of unknown length
	* 12499 51823: contig of 39325 bp in length
	* 51824 51923: gap of unknown length
	* 51924 97334: contig of 45411 bp in length
	* 97335 97434: gap of unknown length
	* 97435 169505: contig of 72071 bp in length
	* 169506 169605: gap of unknown length
	* 169606 232395: contig of 62790 bp in length
	* 232396 232495: gap of unknown length
	* 232496 232603: contig of 108 bp in length.
FEATURES	Location/Qualifiers
source	1..232603
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/chromosome="UNK"
	/clone="RP24-267118"
	1..12398
misc_feature	/note="assembly_name:Contig11"
misc_feature	12499..51823
misc_feature	/note="assembly_name:Contig12"
	51924..97334
	/note="assembly_name:Contig13"
	97435..169505
	/note="assembly_name:Contig14"
	169606..232395
	/note="assembly_name:Contig15"
	232496..232603
	/note="assembly_name:Contig7"
	BASE COUNT 67248 a 46673 c 47059 g 71090 t 533 others
ORIGIN	
Query Match	60.5%; Score 24.8; DB 2; Length 232603;
Best Local Similarity	76.3%; Pred. NO. 94;
Matches 29; Conservative 1; Mismatches 8; Indels 0; Gaps 0;	
QY	4 AGACGCTGGAGGAGGACACCTCTGAGGAGACCGCGGAG 41
	:
Db 124434	ARCTCAGGAGGAGGACACTCTGAGGAGACCGCGGAG 124471
	:
RESULT 9	
BC012882/c	
LOCUS	BC012882 Homo sapiens, clone IMAGE:4153436, mRNA, partial cds.
DEFINITION	BC012882
ACCESSION	BC012882.1 GI:15277573
VERSION	
KEYWORDS	Homo sapiens.
SOURCE	Homo sapiens.
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 612)
AUTHORS	Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (20-AUG-2001) National Institutes of Health, Mammalian
	Gene Collection (MGC), Cancer Genomics Office, National Cancer
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
	USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk
	Email: cgapbs-r@mail.nih.gov
	Tissue Procurement: David N. Louis, M.D.
	CDNA Library Preparation: Life Technologies, Inc.
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
	DNA Sequencing by: Baylor College of Medicine Human Genome
	Sequencing Center
	Center code: BCM-HGSC
	Web site: http://www.hgsc.bcm.tmc.edu/cdna/
	Contact: villaia@bcm.tmc.edu.
	Villaia, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
	A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
	Muzny, D.M., Gibbs, R.A.
	Clone distribution: MGC clone distribution information can be found
	through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
	Series: IRAK Plate: 19 Row: p Column: 4.
FEATURES	Location/Qualifiers
source	1..612
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="IMAGE:4153436"
	/tissue_type="Brain, anaplastic oligodendroglioma with
	1p/19q loss"
	/clone_lib="NCI CGAP_Brn67"
	/lab_host="DH10B"
	/note="Vector: pCMV-SPORT6"
	<1..386
	/codon_start=3
	/product="Unknown (protein for IMAGE:4153436)"
	/protein_id="AAH12882.1"
	/db_xref="GI:15277574"
	/translation="GAFSLGPGLSLHQLKQLRALPALPSLSOLELIDLSNPFHC
	PSARRTPIKRGCGADKVGKKGRL"

```
BASE COUNT      142 a   184 c   178 g   108 t
ORIGIN
Query Match      60.0%; Score 24.6; DB 9; Length 612;
Best Local Similarity 73.2%; Pred. No. 3.7e+02;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGGCGCCGGGAG 41
||||| 1 11 |||||||:||||||| 111 |||
Db 502 AGAAGCCCTGCTGGAGGACGACCCTCAGGTGCCAGGAAG 462
||||| 1 11 |||||||:||||||| 111 |||

RESULT 10
AX078286      2555 bp   DNA      linear      PAT 22-FEB-2001
LOCUS
DEFINITION      Sequence 90 from Patent WO0107471.
ACCESSION      AX078286
VERSION
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 2555)
AUTHORS      Hillman,J.L., Lai,P., Tang,Y.T., Yue,H., Au-Young,J., Bandman,O.,
Azimzai,Y., Yang,J., Lu,D.A., Baughn,M.R., Patterson,C. and Shah,P.
TITLE      Cell cycle and proliferation proteins
JOURNAL      Patent: WO 0107471-A 90 01-FEB-2001;
Incyte Genomics, Inc. (US)
FEATURES
source
location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /note="Incyte ID No: 1868749CB1"
BASE COUNT      577 a   709 c   722 g   547 t
ORIGIN

Query Match      60.0%; Score 24.6; DB 6; Length 2555;
Best Local Similarity 73.2%; Pred. No. 2.8e+02;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGGCGCCGGGAG 41
||||| 1 11 |||||||:||||||| 111 |||
Db 1342 AGAAGCCCTGCTGGAGGACGACCCTCAGGTGCCAGGAAG 1382
||||| 1 11 |||||||:||||||| 111 |||

RESULT 11
AX156560/c
LOCUS
DEFINITION      Sequence 1 from Patent WO0142286.
ACCESSION      AX156560
VERSION
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 2636)
AUTHORS      Khodadoust,M.M.
TITLE      Human leucine rich repeat-containing polypeptide and uses therefor
JOURNAL      Patent: WO 0142286-A 1 14-JUN-2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /note="unnamed protein product"
    /codon_start=1
    /protein_id="CAC42683.1"
    /db_xref="GI:14537548"
    /translation="MLETQMSITRGMKLLVVRGTEKAGVAPVGMGPRSSHTVPP
LVLPDLLVLLLPAPARQAARQPCACICDNRHRRHVACRYQNLTETVPDPAIPELTQRLDI
OQNLLKVI PAAAFQGVPHLTLDLRCEVELVAEGAFRCGLGRLILLINLASNHLRELPO
EALDGLSLRLLELGNALIELRPGTFGALGALATINLAHNALVTLPMAPGQLLRVR
WLRSHNALSVAPEALAGLPALRRLSUHNELQALPGVLSQARGLARLEIGHNPLI
YAGEEDGIALPGIRELLLDGGALQALGPRAFHCPRLHTLDIRGNQDLTLPPLQPGQ
LRLRLQITRCGAARGRPTRVAGAGARALGRVPRRLRGALDALTLPMDLRCPGD
AAQEELERFVAVAGPRAPRPGPRGPGGEERAVAPCPACVCPVPSRHSSCGCGLOA
VPRGFSPTOLLDRNRHFPSPVRAAFPGLGLHLSHLQHOGIALEAGALAGLRLLI
YLTLSDNOLAGLSAAALAGVPRGLGYLYLERNRFLQVPPGAALRALRSLFSLHODNAV
RLAPDGLRTRALRWYLSGNRITVSLGALGPARELEKHLDRNQLREVPTGALLEG
PALLELQUSGNPLRALRDGAFQVGRSLQHLFLNSSGLEIICPGAFSGIGPLGSLHL
QKNQRLALPALPSQLLELIDLSNPFHDCQLLPLRHLWTGLNLRVGATCATPPNAR
GQRVKAAAADFEDCPGWAARKRTPIPASRPSARRTPIKRGCGGADKVGKEGCL"
BASE COUNT      403 a   922 c   876 g   431 t
ORIGIN

Query Match      60.0%; Score 24.6; DB 6; Length 2636;
Best Local Similarity 73.2%; Pred. No. 2.7e+02;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGGCGCCGGGAG 41
||||| 1 11 |||||||:||||||| 111 |||
Db 2533 AGAAGCCCTGCTGGAGGACGACCCTCAGGTGCCAGGAAG 2493
||||| 1 11 |||||||:||||||| 111 |||

RESULT 12
HSN801538
LOCUS
DEFINITION      Homo sapiens mRNA; cDNA DKFzp7611141 (from clone DKFp7611141);
complete cds.
ACCESSION      HSN801538
VERSION
KEYWORDS
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 3071)
AUTHORS      Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE      Direct Submission
JOURNAL      Submitted (12-MAR-2002) MTPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT      Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKFz7611141) is available at the RZPD in Berlin. Please contact
the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
source
location/Qualifiers
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /map="22q13.31-13.33"
    /clone="DKFzp7611141"
    /tissue_type="amygdala"
    /clone_lib="761 (synonym: hamy2). Vector pSport1; host
DH10B; sites NotI + SalI"
    /dev_stage="adult"
    /gene="DKFzp7611141"
    /size=1908
    /gene="DKFzp7611141"
    /note="similarity to Human l(3)mbt protein homolog mRNA"
    /codon_start=1
    /product="hypothetical protein"
    /protein_id="CAC66499.1"
    /db_xref="GI:13276635"
    /db_xref="SPTREMBL:Q9BQ12"
    /translation="MEKPRSIETPSEPMEEEDDDLEFGGYDIFRSYNSVSGSES
SSYLESSEAEENEDREAGELPTSPHLILSPTRSLDGGSGSEPAVCMCGIVGTREAF
OQNLLKVI PAAAFQGVPHLTLDLRCEVELVAEGAFRCGLGRLILLINLASNHLRELPO
EALDGLSLRLLELGNALIELRPGTFGALGALATINLAHNALVTLPMAPGQLLRVR
WLRSHNALSVAPEALAGLPALRRLSUHNELQALPGVLSQARGLARLEIGHNPLI
YAGEEDGIALPGIRELLLDGGALQALGPRAFHCPRLHTLDIRGNQDLTLPPLQPGQ
LRLRLQITRCGAARGRPTRVAGAGARALGRVPRRLRGALDALTLPMDLRCPGD
AAQEELERFVAVAGPRAPRPGPRGPGGEERAVAPCPACVCPVPSRHSSCGCGLOA
VPRGFSPTOLLDRNRHFPSPVRAAFPGLGLHLSHLQHOGIALEAGALAGLRLLI
YLTLSDNOLAGLSAAALAGVPRGLGYLYLERNRFLQVPPGAALRALRSLFSLHODNAV
RLAPDGLRTRALRWYLSGNRITVSLGALGPARELEKHLDRNQLREVPTGALLEG
PALLELQUSGNPLRALRDGAFQVGRSLQHLFLNSSGLEIICPGAFSGIGPLGSLHL
QKNQRLALPALPSQLLELIDLSNPFHDCQLLPLRHLWTGLNLRVGATCATPPNAR
GQRVKAAAADFEDCPGWAARKRTPIPASRPSARRTPIKRGCGGADKVGKEGCL"
BASE COUNT      403 a   922 c   876 g   431 t
ORIGIN
```

FSKTRFCSVCSRSYSSNSKKASILARLOGKPPYKAKVLHKAWSAKIGAFHLSQG  
TQGLADGTPGQDALVIGDFWKGFLKDHYSYKAAPVSCFKHVLPLYDQWEDVMGKVEV  
LNSDAVPSRVWIASVIQTAGYRVLLRYEGFENDASHDFWNLGTVDVHPIGWCAIN  
SKILVPRTHAKFTDQWGLKMLKRVGSRTPVDFHIKVMESKYPFROGMRLEVVDK  
SOVSRTMAVVTIGRULRYLEDGSDDDFWCHWMSPLIHPVGSRRVGHGKIMSE  
RSDMAHPTFKIKYCDVAPYLFKVRVYTEGGWFEEGKMLEADPLNLGNICVATV  
CKVLLDGYLMICVDGSPDGLDFWCFYHASSHAIFATPCQKNDILTPPKGYEATTF  
NWNLEKTKSAAAPFNFMDCPNHGFVKMKLEAVDLMEPRLICVATVKRVVHRL  
SHFGDWDSEYDQWDCSPDIYPVGVWCELTGYQLQPPVAAGVSGRPRKL"  
3052  
polyA\_site  
/gene-"DKF2p761I141"  
BASE COUNT 666 a 882 c 858 g 665 t  
ORIGIN

Query Match 60.0%; Score 24.6; DB 9; Length 3071;  
Best Local Similarity 73.2%; Pred. No. 2.7e+02;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGAAGCAGTGGAGGACRACCTCAGCAGCCCCGGGAG 41  
||||| 1 || |||||:||||| ||| || ||  
Db 2471 AGAAGCCCTGCTGGAGGACGCCCTCAGGGTCCAGGAAG 2511

RESULT 13  
AK056443  
LOCUS  
DEFINITION Homo sapiens CDNA FLJ31881 fis, clone NT2RP7002829, weakly similar to Scm-related gene containing four mbt domains.  
ACCESSION AK056443  
VERSION AK056443.1 GI:16551846  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens teratocarcinoma cell\_line:NT2 cDNA to mRNA,  
clone\_lib:NT2RP7 clone:NT2RP7002829.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiya,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saico,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuna,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 3188)  
Isogai,T., Otsuki,T. and Sugiyama,T.  
Direct Submission  
TITLE  
JOURNAL  
AUTHORS  
TITLE  
JOURNAL

COMMENT  
E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES  
source  
1. .3188  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="NT2RP7002829"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2RP7"  
/note="cloning vector: pME18SFL3-mRNA from NT2 neuronal precursor cells after 5-weeks retinoic acid (RA) induction."  
BASE COUNT 725 a 869 c 909 g 685 t

ORIGIN  
Query Match 60.0%; Score 24.6; DB 9; Length 3188;  
Best Local Similarity 73.2%; Pred. No. 2.6e+02;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 1 AGAAGCAGTGGAGGACRACCTCAGCAGCCCCGGGAG 41  
||||| 1 || |||||:||||| ||| || ||  
Db 1988 AGAAGCCCTGCTGGAGGACGCCCTCAGGGTCCAGGAAG 2028

RESULT 14  
HSA305226  
LOCUS  
DEFINITION Homo sapiens mRNA for H-1(3)mbt-like protein, alternative variant a.  
ACCESSION AJ305226  
VERSION AJ305226.1 GI:13940238  
KEYWORDS alternative splicing; H-1(3)mbt-like gene; H-1(3)mbt-like protein.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1  
AUTHORS Wismar,J.  
TITLE Molecular characterization of h-1(3)mbt-like: a new member of the human mbt family  
JOURNAL FEBS Lett. 507 (1), 119-121 (2001)  
MEDLINE 21538645  
PUBMED 11682070  
REFERENCE 2 (bases 1 to 3194)  
AUTHORS Wismar,J.  
TITLE Direct Submission  
JOURNAL Submitted (25-JAN-2001) Wismar J., Johannes Gutenberg Universitaet, Institut fuer Genetik, Becherweg 32, 55099 Mainz, GERMANY  
FEATURES  
source  
1. .3194  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosomes="22"  
/map="22q13.31-33"  
40. .2157  
/gene="H-1(3)mbt-like"  
40. .2157  
/gene="H-1(3)mbt-like"  
/note="alternative variant a"  
/codon\_start=1  
/product="H-1(3)mbt-like protein"  
/protein\_id="CAC37794.1"  
/db\_xref="GI:13940239"

gene  
CDS  
/translation="MEKPRSIETPSSEPMEEEDDDLELFGYDSFRSYNSVSGSES  
SSVLEPSSAENEDREAGELPTSLHLLSPGTPRSLDGSGSEPAVTEMGIVGTREAF  
FSKTRFCSVCSRSYSSNSKKASILARLOGKPPYKAKVLHKAWSAKIGAFHLSQG  
TQGLADGTPGQDALVIGDFWKGFLKDHYSYKAAPVSCFKHVLPLYDQWEDVMGKVEV  
LNSDAVPSRVWIASVIQTAGYRVLLRYEGFENDASHDFWNLGTVDVHPIGWCAIN  
SKILVPRTHAKFTDQWGLKMLKRVGSRTPVDFHIKVMESKYPFROGMRLEVVDK  
SOVSRTMAVVTIGRULRYLEDGSDDDFWCHWMSPLIHPVGSRRVGHGKIMSE  
RSDMAHPTFKIKYCDVAPYLFKVRVYTEGGWFEEGKMLEADPLNLGNICVATV  
CKVLLDGYLMICVDGSPDGLDFWCFYHASSHAIFATPCQKNDILTPPKGYEATTF  
NWNLEKTKSAAAPFNFMDCPNHGFVKMKLEAVDLMEPRLICVATVKRVVHRL  
SHFGDWDSEYDQWDCSPDIYPVGVWCELTGYQLQPPVAAPATPKAKEATFKKKLL  
QFGKKRIPPTKTRPLRQSGKKPLLEDDPQGARKISSEPPVGEIIVRVRKEHLDVA  
SPDKASSPELPVSVENIKETDD"  
BASE COUNT 738 a 869 c 906 g 681 t  
ORIGIN

Query Match 60.0%; Score 24.6; DB 9; Length 3194;  
Best Local Similarity 73.2%; Pred. No. 2.6e+02;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 1 AGAAGCAGTGGAGGACRACCTCAGCAGCCCCGGGAG 41  
||||| 1 || |||||:||||| ||| || ||  
Db 1976 AGAAGCCCTGCTGGAGGACGCCCTCAGGGTCCAGGAAG 2016

RESULT 15  
BC017191  
LOCUS BC017191  
DEFINITION Homo sapiens, hypothetical protein DKFZp7611141, clone MGC:2476  
IMAGE:3138444, mRNA, complete cds.  
ACCESSION BC017191  
VERSION BC017191.1 GI:16877934  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 3222)  
Strausberg, R.  
Direct Submission  
Submitted (05-NOV-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nhgri.nih.gov](mailto:nisc.mgc@nhgri.nih.gov)  
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,  
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,  
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 4 Row: D Column: 5  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 13899349.

## FEATURES

source  
1..3222  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="LocusID:83746"  
/db\_xref="taxon:9606"  
/clone="MGC:2476 IMAGE:3138444"  
/tissue\_type="placenta, choriocarcinoma"  
/clone\_lib="NIH\_MGC\_21"  
/lab\_host="DH10B-R"  
/note="Vector: pORF7"  
53..2170  
/codon\_start=1  
/product="hypothetical protein DKFZp7611141"  
/protein\_id="AAH17191.1"  
/db\_xref="GI:16877935"  
/translation="MEKRSIETPSPSEPMEEEDDDLELFGYDPSFRSYNSVGSSES  
SYLSESEAEDEAGELPTPLHLSPGTRSLDGSSEPAVCENCGIVGTREAF  
FSKTRFSRCSYSSNSKASILARLQGPPTKAKVLHKAWSAKIGAFHLSQG  
TQLADGPTQGDALVLDGKGLKSHSKAAPVSCFKHPLYDOWEDVNMKMKVEV  
LNSDAVLSRVYIASVTAGYRLLRVGEFENDASHDFWNLGTVDPVHPIGHCAIN  
SKILVPRTHAKFTDVKGLMKLVGSRITLPVDEHIKVESMKYPPQGMRLVVDK  
SOVSTRNAVDTVLGGRLRLLEDGSDDDFWCHWMSPLIHPVGSRRRVGHGIRMSE  
RRSDNAHPPTFKITCDAPVFLFKRVAVYTEGGWFEGMKLEADPLNLGNICVATY  
CKVLDDGLMTCVDGPTDGLDFWCFYHASSHAIFPATFCQKNDIELTTPPKGYEATF  
NWNELYTKSKAAPSRLFNMDCPNHGFKVGMKLEAVDLMEPRLICVATVKRVYHRL  
SIHFDGSDSEYDQWDCSPDIYPVWGCELTGYOLOPPVAEPATPLKAKETKKKK  
QFGKKRKRPPTKTPPLRQSGKPLLEDPOGARKISSEPVPGELIIVRVKEHLDVA  
SPDKASSPELPVSVENIKQETDD"

## CDS

BASE COUNT 756 a 870 c 913 g 683 t  
ORIGIN  
Query Match 60.0%; Score 24.6; DB 9; Length 3222;  
Best Local Similarity 73.2%; Pred. No. 2.6e+02;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 1 AGAAGCCAGTGGAGGAGGACRACCCCTCAGGCAGCCCGGGAG 41  
||||| | || |||||:||||||| ||| || ||  
Db 1989 AGAAGCCCTGCTGGAGGAGCACCCTCAGGGTGCAGGAAG 2029

Search completed: February 11, 2003, 05:31:29  
Job time : 196.088 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 21:06:50 ; Search time 141.088 Seconds  
(without alignments)  
8457.264 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_860\_900

Perfect score: 41

Sequence: 1 gttgtgagagagaatgtgtgc.....ctaagtgtcagtgtagttct 41

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vi:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vi:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rod:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_htgo\_hum:\*
- 40: em\_htgo\_mus:\*
- 41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	40.6	99.0	1669	6	AX207224	AX207224 Sequence
2	40.6	99.0	1680	6	AX394457	AX394457 Sequence
3	40.6	99.0	5503	9	HUMCYP2DG	M33189 Human debri
4	40.6	99.0	9432	6	AX394456	AX394456 Sequence
5	40.6	99.0	9432	9	HUMCYP2D6	M33388 Human cyloc
6	40.6	99.0	13677	9	HSCYP2D7B	X58468 Human CYP2D
7	39	95.1	13278	9	HSCYP2D7A	X58467 Human CYP2D
8	39	95.1	17060	9	HUMCYP8P	M33387 Human debri
9	39	95.1	114846	9	HS257120	AL021878 Human DNA
10	32.6	79.5	5884	6	AX345458	AX345458 Sequence
11	32.6	79.5	5884	6	AX348344	AX348344 Sequence
12	27.8	67.8	207709	2	AC087799	AC087799 Mus muscu
13	27.4	66.8	191496	2	AL772168	AL772168 Mus muscu
14	27.4	66.8	220770	2	AL772150	AL772150 Mus muscu
15	26.4	64.4	160365	2	AC123144	AC123144 Rattus no
16	26.2	63.9	65741	2	AC119833	AC119833 Mus muscu
17	26.2	63.9	72677	2	AC123691	AC123691 Mus muscu
18	26.2	63.9	94409	2	AC096217	AC096217 Rattus no
19	26.2	63.9	188142	2	AC126883	AC126883 Rattus no
20	26.2	63.9	220166	2	AC022781	AC022781 Mus muscu
21	26.2	63.9	220176	10	AL591065	AL591065 Mouse DNA
22	25.8	62.9	154888	2	AC114830	AC114830 Rattus no
23	25.8	62.9	178746	10	AC087067	AC087067 Rattus no
24	25.8	62.9	190418	2	AC118370	AC118370 Rattus no
25	25.4	62.0	131364	2	AC128769	AC128769 Rattus no
26	25.2	61.5	53263	2	AC018576	AC018576 Homo sapi
27	25.2	61.5	132120	9	AC026733	AC026733 Homo sapi
28	25.2	61.5	150468	2	AC118858	AC118858 Rattus no
29	25.2	61.5	157331	2	AC015774	AC015774 Homo sapi
30	25.2	61.5	171070	2	AC012288	AC012288 Homo sapi
31	25.2	61.5	180219	9	AC099520	AC099520 Homo sapi
32	25.2	61.5	259272	2	AC096436	AC096436 Rattus no
33	25	61.0	143096	9	HS140C12	AL008628 Homo sapi
34	25	61.0	147153	2	AC123445	AC123445 Rattus no
35	25	61.0	173713	2	AC095840	AC095840 Rattus no
36	25	61.0	174096	2	AC097177	AC097177 Rattus no
37	25	61.0	185967	9	AC012508	AC012508 Homo sapi
38	25	61.0	201275	10	AL603662	AL603662 Mouse DNA
39	24.8	60.5	137127	2	AC113343	AC113343 Homo sapi
40	24.8	60.5	172493	2	AC109559	AC109559 Rattus no
41	24.8	60.5	174418	2	AC122098	AC122098 Rattus no
42	24.8	60.5	174940	2	AC106570	AC106570 Rattus no
43	24.8	60.5	179526	2	AC034133	AC034133 Homo sapi
44	24.8	60.5	180331	9	AC007375	AC007375 Homo sapi
45	24.8	60.5	180727	2	AC027221	AC027221 Homo sapi

ALIGNMENTS

RESULT 1	AX207224	AX207224	Sequence 1	1669 bp	DNA	linear	PAT 30-AUG-2001
AX207224	LOCUS	AX207224	Sequence 1 from Patent WO0155432.				
	DEFINITION	AX207224					
	ACCESSION	AX207224					
	VERSION	AX207224.1	GI:15394976				
	KEYWORDS						
	SOURCE		Synthetic construct.				
	ORGANISM		artificial construct				
	REFERENCE		1 (bases 1 to 1669)				
	AUTHORS		Raimundo, S. and Zanger, U.				
	TITLE		Polymorphisms in the human cyp2d6 gene promoter region and their use in diagnostic and therapeutic applications				
	JOURNAL		Patent: WO 0155432-A 1 02-AUG-2001;				







```

PUBMED      1358797
COMMENT      See X58468, and Am. J. Hum. Genet. 47:994-1001(1990)).
FEATURES
source
1..13278
    Location/Qualifiers
        organism="Homo sapiens"
        db_xref="taxon:9606"
        chromosome="22"
        clone="42"
        clone_lib="EMBL4"
        dev_stage="adult"
        1154..5489
            /gene="CYP2D7AP"
        join(1154..1424,2125..2296,2823..2975,3064..3225,
3651..3827,4020..4161,4356..4542,4998..5139,5238..5489)
            /gene="CYP2D7AP"
        1154..1424
            /gene="CYP2D7AP"
        number=1
        join(1242..1424,2125..2296,2823..2975,3064..3225,
3651..3827,4020..4161,4356..4542,4998..5139,5238..5416)
            /gene="CYP2D7AP"
        pseudo
        /codon_start=1
        /db_xref="PID:e34292"
        1425..2124
            /gene="CYP2D7AP"
        number=1
        2125..2296
            /gene="CYP2D7AP"
        number=2
        /gene="CYP2D7AP"
        2297..2822
            /gene="CYP2D7AP"
        number=2
        2823..2975
            /gene="CYP2D7AP"
        number=3
        2976..3063
            /gene="CYP2D7AP"
        number=3
        3064..3225
            /gene="CYP2D7AP"
        number=4
        3226..3650
            /gene="CYP2D7AP"
        number=4
        3651..3827
            /gene="CYP2D7AP"
        number=5
        3828..4019
            /gene="CYP2D7AP"
        number=5
        4020..4161
            /gene="CYP2D7AP"
        number=6
        4162..4355
            /gene="CYP2D7AP"
        number=6
        4356..4542
            /gene="CYP2D7AP"
        number=7
        4543..4997
            /gene="CYP2D7AP"
        number=7
        4998..5139
            /gene="CYP2D7AP"
        number=8
        5140..5237
            /gene="CYP2D7AP"
        number=8
        5238..5489
            /gene="CYP2D7AP"
        number=9
        7887..7908
            /gene="CYP2D7AP"
        repeat_region

BASE COUNT      2902 a      3664 c      3968 g      2744 t
ORIGIN
Query Match      95.1%; Score 39; DB 9; Length 13278;
Best Local Similarity 95.1%; Pred. NO. 2c-05;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY      1 GTGTGAGAGAGAAATGTGTGTCYCTAAGTGTCAAGTGTGAGTCT 41
        |||||||||||||||||||:|||||||||||||||||
Db      482 GTGTGAGAGAGAAATGTGTGCCCCGTGAGTGTCAAGTGTGAGTCT 522

RESULT 8
HUMCYP8P
LOCUS      HUMCYP8P      17060 bp      DNA      linear      PRI 09-NOV-1994
DEFINITION      Human debrisoquine 4-hydroxylase (CYP2D8P) and (CYP2D7) pseudogenes
complete sequences.
ACCESSION      M33387
VERSION      M33387.1 GI:181320
KEYWORDS      debrisoquine 4-hydroxylase.
SOURCE      Human DNA, clones lambda-2D-A and lambda-2D-B.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 17060)
AUTHORS      Kimura,S., Umeno,M., Skoda,R.C., Meyer,U.A. and Gonzalez,F.J.
TITLE      The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and
identification of the polymorphic CYP2D6 gene, a related gene, and
a pseudogene
JOURNAL      Am. J. Hum. Genet. 45 (6), 889-904 (1989)
MEDLINE      90072069
PUBMED      2574001
COMMENT      Draft entry and computer-readable sequence for [1] kindly submitted
        by S. Kimura, 29-MAR-1990.
FEATURES
source
1..17060
    Location/Qualifiers
        TATA_signal      1276..1282
        prim_transcript      1304..6570
            /note="CYP2D8P mRNA and introns"
            join(1392..1568,3189..3360,3907..4059,4148..4310,
4758..4934,5121..5262,5467..5651,6101..6242,6339..6516)
                /gene="CYP2D8P"
            join(1392..1568,3189..3360,3907..4059,4148..4310,
4758..4934,5121..5262,5467..5651,6101..6242,6339..6516)
                /gene="CYP2D8P"
            pseudo
            /codon_start=1
            /product="debrisoquine 4-hydroxylase"
            <1392..1568
                /gene="CYP2D8P"
            /number=1
            3189..3360
                /gene="CYP2D8P"
            /number=2
            pseudo
            3907..4059
                /gene="CYP2D8P"
            /number=3
            pseudo
            4148..4310
                /gene="CYP2D8P"
            /number=4
            pseudo
            4758..4934
                /gene="CYP2D8P"
            /number=5
            pseudo
            5121..5262
                /gene="CYP2D8P"
            /number=6
```



repeat\_region /note="match: GSS: Em:AQ124532"  
13208..114275  
/note="MER11C repeat: matches 1..1057 of consensus"  
repeat\_region 16086..16507  
/note="L1MD repeat: matches 1552..1964 of consensus"  
repeat\_region 17055..17127  
/note="L1MD repeat: matches 1482..1552 of consensus"  
repeat\_region 17434..17688  
/note="L1MD repeat: matches 1222..1482 of consensus"  
19152..19153  
/note="clone RP1-18601  
aa in this entry  
deletion"  
/replace="aaaata"  
repeat\_region 19896..19939  
/note="L1P repeat: matches 2416..2459 of consensus"  
repeat\_region 20246..20284  
/note="L1P repeat: matches 2378..2416 of consensus"  
repeat\_region 20594..20660  
/note="L1P repeat: matches 2311..2378 of consensus"  
repeat\_region 20965..20982  
/note="L1P repeat: matches 2294..2311 of consensus"  
repeat\_region 21270..22130  
/note="L1P repeat: matches 1445..2294 of consensus"  
repeat\_region 22425..22499  
/note="L1P repeat: matches 1370..1445 of consensus"  
repeat\_region 22502..22994  
/note="LTR10B repeat: matches 1..510 of consensus"  
repeat\_region 23020..25379  
/note="BaEV-int repeat: matches 4243..6625 of consensus"  
25398..25662  
/note="match: STS: Em:G07321"  
25597..26083  
/note="match: GSS: Em:B13982"  
25611..25853  
/note="match: GSS: Em:B14069"  
complement(26289..26446)  
/note="match: GSS: Em:AF046780"  
26890..26995  
/note="match: GSS: Em:A2083430"  
27147..28076  
/note="HERVI repeat: matches 4455..5408 of consensus"  
repeat\_region 28386..30853  
/note="HERVI repeat: matches 2098..4455 of consensus"  
repeat\_region 30843..30912  
/note="HERVI repeat: matches 2098..2163 of consensus"  
repeat\_region 30902..30971  
/note="HERVI repeat: matches 2098..2163 of consensus"  
repeat\_region 30961..31356  
/note="HERVI repeat: matches 1561..2163 of consensus"  
repeat\_region 31873..33456  
/note="HERVI repeat: matches 9..1561 of consensus"  
repeat\_region 33464..33959  
/note="LTR10B repeat: matches 1..510 of consensus"  
repeat\_region 33960..34110  
/note="L1Pa repeat: matches 1226..1378 of consensus"  
repeat\_region 34415..35367  
/note="L1Pa repeat: matches -656..1226 of consensus"  
repeat\_region 37163..37323  
/note="Harlequin repeat: matches 912..1086 of consensus"  
repeat\_region 37324..37437  
/note="57 copies 2 mer ga 79% conserved"  
complement(join(42379..42591,42916..43112))  
/note="match: STS: Em:G27508"  
misc\_feature complement(join(42491..42603,42916..42945))  
/note="match: STS: Em:G43129"  
43992..44915  
/note="PTR5 repeat: matches 40..93 of consensus"  
repeat\_region 44916..45564  
/note="LTR12 repeat: matches 5..671 of consensus"  
misc\_feature 48350..49467  
/note="CpG island"  
/evidence=not\_experimental

repeat\_region 54417..54450  
/note="17 copies 2 mer ca 100% conserved"  
54578..55083  
/note="MER77 repeat: matches 30..587 of consensus"  
58051..59086  
/note="CpG island"  
/evidence=not\_experimental  
65436..65550  
/note="match: GSS: Em:AQ035975"  
65916..65961  
/note="23 copies 2 mer tt 82% conserved"  
complement(65929..66250)  
/note="match: STS: Em:G27630"  
66514..66549  
/note="18 copies 2 mer aa 80% conserved"  
66530..66802  
/note="match: GSS: Em:AQ617313"  
66577..66930  
/note="match: GSS: Em:AQ084812"  
66603..66893  
/note="match: GSS: Em:AQ044403"  
66974..67017  
/note="22 copies 2 mer cc 75% conserved"  
complement(69196..69374)  
/note="match: GSS: Em:B14383"  
69710..70103  
/note="match: GSS: Em:AQ181535"  
complement(78378..78961)  
/note="match: GSS: Em:AQ389013"  
complement(78421..78533)  
/note="match: GSS: Em:AQ042556"  
complement(78451..78977)  
/note="match: GSS: Em:AQ533248"  
complement(81846..82274)  
/note="match: GSS: Em:AQ618257"  
82587..83003  
/note="L1MA5A repeat: matches 5857..6292 of consensus"  
83609..83702  
/note="MER21B repeat: matches 703..787 of consensus"  
83697..83797  
/note="MER21B repeat: matches 686..794 of consensus"  
repeat\_region 83910..84357  
/note="MER21B repeat: matches 76..540 of consensus"  
85375..85430  
/note="28 copies 2 mer aa 76% conserved"  
complement(90629..90900)  
/note="match: GSS: Em:B13983"  
96343..96384  
/note="11 copies 2 mer ac 100% conserved"  
99274..99444  
/note="Other . Weak data"  
104028..104091  
/note="32 copies 2 mer aa 67% conserved"  
complement(107923..108378)  
/note="match: GSS: Em:AQ572846"  
  
Query Match 95.1%; Score 39; DB 9; Length 114846;  
Best Local Similarity 95.1%; Pred. No. 1.5e-05;  
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GTGTGAGAGAGATGTGTCTTAAGTGTGAGTGTGAGTGT 41  
|||||  
DB 51113 GTGTGAGAGAGATGTGTCTTAAGTGTGAGTGTGAGTGT 51073  
  
RESULT 10  
AX345458  
LOCUS AX345458 5884 bp DNA linear PAT 01-FEB-2002  
DEFINITION Sequence 529 from Patent WO0200928.  
ACCESSION AX345458  
VERSION AX345458.1 GI:18493344  
KEYWORDS  
SOURCE synthetic construct.

```

ORGANISM      synthetic construct
               artificial sequences.
REFERENCE
1
AUTHORS      Olek,A., Piepenbrock,C. and Berlin,K.
TITLE        Diagnosis of diseases associated with the immune system
JOURNAL      Patent: WO 0200928-A 529 03-JAN-2002;
               Epigenomics AG (DE)
FEATURES
source
1..5884
   /organism="synthetic construct"
   /db_xref="taxon:32630"
BASE COUNT    1259 a 92 c 1514 g 3019 t
ORIGIN
Query Match      79.5%; Score 32.6; DB 6; Length 5884;
Best Local Similarity 85.4%; Pred. No. 0.0088;
Matches 35; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGAAATGTGTGCYCTAAGTGTCAAGTGTGAGTCT 41
|||||
Db 4241 GTGTGAGAGAGAAATGTGTGTTTTCAGTGTGAGTGTGAGTTT 4281

RESULT 11
AX348344
LOCUS
DEFINITION      Sequence 39 from Patent WO0202806.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequences.
1
REFERENCE
AUTHORS      Olek,A., Piepenbrock,C. and Berlin,K.
TITLE        Method and nucleic acids for pharmacogenomic methylation analysis
JOURNAL      Patent: WO 0202806-A 39 10-JAN-2002;
               Epigenomics AG (DE)
FEATURES
source
1..5884
   /organism="synthetic construct"
   /db_xref="taxon:32630"
   /note="Chemically treated genomic DNA (Homo sapiens)"
BASE COUNT    1259 a 92 c 1514 g 3019 t
ORIGIN
Query Match      79.5%; Score 32.6; DB 6; Length 5884;
Best Local Similarity 85.4%; Pred. No. 0.0088;
Matches 35; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGAAATGTGTGCYCTAAGTGTCAAGTGTGAGTCT 41
|||||
Db 4241 GTGTGAGAGAGAAATGTGTGTTTTCAGTGTGAGTGTGAGTTT 4281

RESULT 12
AC087799/c
LOCUS
DEFINITION      Mus musculus chromosome 19 clone rp23-198m10 strain C57BL/6J,
               WORKING DRAFT SEQUENCE, 2 ordered pieces.
ACCESSION
VERSION
KEYWORDS      HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 207709)
AUTHORS      Loh,P., Qi,S., Ford,B., Hine,R. and Roe,B.A.
TITLE        Mus musculus Chromosome 16 BAC Clone rp23-198m10
JOURNAL      Unpublished
REFERENCE
2 (bases 1 to 207709)

```

---

```

AUTHORS      Loh,P., Qi,S., Ford,B., Hine,R. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (26-JAN-2001) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
REFERENCE
3 (bases 1 to 207709)
AUTHORS      Loh,P., Qi,S., Ford,B., Hine,R. and Roe,B.A.
TITLE        Direct Submission
JOURNAL      Submitted (01-AUG-2002) Department Of Chemistry And Biochemistry,
               The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
               OK 73019, USA
COMMENT
On Jul 12, 2002 this sequence version replaced gi:21672179.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 60622: contig of 60622 bp in length
* 60623 60722: gap of unknown length
* 60723 207709: contig of 146987 bp in length.
FEATURES
source
1..207709
   /organism="Mus musculus"
   /strain="C57BL/6J"
   /db_xref="taxon:10090"
   /chromosome="19"
   /clone="rp23-198m10"
   /clone_lib="RPC1 - 23 Female (C57BL/6J) Mouse BAC Library"
   /note="This is one of two clones in well rp23-198m10"
BASE COUNT    56334 a 46940 c 45622 g 58713 t 100 others
ORIGIN
Query Match      67.8%; Score 27.8; DB 2; Length 207709;
Best Local Similarity 78.0%; Pred. No. 0.48;
Matches 32; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGAAATGTGTGCTAAGTGTCAAGTGTGAGTCT 41
|||||
Db 20108 GGGAGAGAGAGAAATGTGTCTCTGAGTGTGTGTGTGTGT 20068

RESULT 13
AL772168
LOCUS
DEFINITION      Mus musculus chromosome 4 clone RP23-419B9, *** SEQUENCING IN
               PROGRESS ***, 13 unordered pieces.
ACCESSION
VERSION
KEYWORDS      HTG; HTGS_PHASE1.
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 191496)
AUTHORS      Plumb,B.
TITLE        Direct Submission
JOURNAL      Submitted (17-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
               Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
               humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
               ----- Genome Center
               Center: Wellcome Trust Sanger Institute
               Center code: SC
               Web site: http://www.sanger.ac.uk
               Contact: humquery@sanger.ac.uk

```



----- Project Information  
Center project name: bm419B9  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 188691 bases at least Q40  
Consensus quality: 189494 bases at least Q30  
Consensus quality: 189925 bases at least Q20  
Insert size: 190296; sum-of-contigs  
Insert size: 193147; 5.6% error; agarose-fp  
Quality coverage: 5.82x in Q20 bases; sum-of-contigs Quality  
coverage: 6.08x in Q20 bases; agarose-fp  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 13 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 5134: contig of 5134 bp in length  
\* 5135 5234: gap of 100 bp  
\* 5235 13968: contig of 8734 bp in length  
\* 13969 14068: gap of 100 bp  
\* 14069 22628: contig of 8560 bp in length  
\* 22629 22728: gap of 100 bp  
\* 22729 77186: contig of 54458 bp in length  
\* 77187 77286: gap of 100 bp  
\* 77287 101492: contig of 24206 bp in length  
\* 101493 101592: gap of 100 bp  
\* 101593 104529: contig of 2937 bp in length  
\* 104530 104629: gap of 100 bp  
\* 104630 112436: contig of 7807 bp in length  
\* 112437 112536: gap of 100 bp  
\* 112537 123477: contig of 10941 bp in length  
\* 123478 123577: gap of 100 bp  
\* 123578 134814: contig of 11237 bp in length  
\* 134815 134914: gap of 100 bp  
\* 134915 152680: contig of 17766 bp in length  
\* 152681 152780: gap of 100 bp  
\* 152781 167109: contig of 14329 bp in length  
\* 167110 167209: gap of 100 bp  
\* 167210 186640: contig of 19431 bp in length  
\* 186641 186740: gap of 100 bp  
\* 186741 191496: contig of 4756 bp in length.

FEATURES  
source

/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosomes="4"  
/clone="RP23-419B9"  
/clone\_lib="RPCI-23"  
/..5134  
/note="assembly\_fragment:00091  
fragment\_chain:1  
clone\_end:SP6  
vector\_side:left"  
5235..13968  
/note="assembly\_fragment:02295  
fragment\_chain:1"  
14069..22628  
/note="assembly\_fragment:01955  
fragment\_chain:1"  
22729..77186  
/note="assembly\_fragment:01216  
fragment\_chain:1"  
77287..101492  
/note="assembly\_fragment:00897  
fragment\_chain:1"  
101593..104529  
/note="assembly\_fragment:02326  
fragment\_chain:1"

misc\_feature 104630..112436  
/note="assembly\_fragment:00083  
fragment\_chain:1"  
112537..123477  
/note="assembly\_fragment:01225  
fragment\_chain:1"  
123578..134814  
/note="assembly\_fragment:00684  
fragment\_chain:1"  
134915..152680  
/note="assembly\_fragment:01654  
fragment\_chain:1"  
152781..167109  
/note="assembly\_fragment:00069  
fragment\_chain:1"  
167210..186640  
/note="assembly\_fragment:00992  
fragment\_chain:1"  
186741..191496  
/note="assembly\_fragment:01008  
fragment\_chain:1  
clone\_end:T7  
vector\_side:right"  
BASE COUNT 54527 a 42820 c 41977 g 50956 t 1216 others  
ORIGIN  
Query Match 66.8%; Score 27.4; DB 2; Length 191496;  
Best Local Similarity 79.5%; Pred.No.0.7;  
Matches 31; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 1 GTCGAGAGAGAATGTGTCYCTAAGTGTGTCAGTGTGAGT 39  
||| ||||| ||||| ||| ||||| ||||| ||  
Db 177876 GTCGAGAGAGAATGTGTCGTCGTCGTCGAGTGTGT 177914  
RESULT 14  
AL772150/c  
LOCUS AL772150 220770 bp DNA linear HTG 09-AUG-2002  
DEFINITION Mus musculus chromosome 4 clone RP23-457P12, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 20 unordered pieces.  
ACCESSION AL772150  
VERSION AL772150.3 GI:22204554  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 220770)  
AUTHORS Ellington,A.  
TITLE Direct Submission  
JOURNAL Submitted (03-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Aug 11, 2002 this sequence version replaced gi:21911661.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: bm457P12  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 214916 bases at least Q40  
Consensus quality: 216904 bases at least Q30  
Consensus quality: 217911 bases at least Q20  
Insert size: 218870; sum-of-contigs  
Insert size: 180924; 8.4% error; agarose-fp  
Quality coverage: 5.65x in Q20 bases; sum-of-contigs Quality  
coverage: 8.07x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 20 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 5304: contig of 5304 bp in length  
\* 5305 5404: gap of 100 bp  
\* 5405 50486: contig of 45082 bp in length  
\* 50487 50586: gap of 100 bp  
\* 50587 62563: contig of 11977 bp in length  
\* 62564 62663: gap of 100 bp  
\* 62664 77675: contig of 15012 bp in length  
\* 77676 77775: gap of 100 bp  
\* 77776 79885: contig of 2110 bp in length  
\* 79886 79985: gap of 100 bp  
\* 79886 82568: contig of 2583 bp in length  
\* 82569 82668: gap of 100 bp  
\* 82669 85961: contig of 3293 bp in length  
\* 85962 86061: gap of 100 bp  
\* 86062 88312: contig of 2251 bp in length  
\* 88313 88412: gap of 100 bp  
\* 88413 92031: contig of 3619 bp in length  
\* 92032 92131: gap of 100 bp  
\* 92132 95041: contig of 2910 bp in length  
\* 95042 95141: gap of 100 bp  
\* 95142 98148: contig of 3007 bp in length  
\* 98149 98248: gap of 100 bp  
\* 98249 100384: contig of 2136 bp in length  
\* 100385 100484: gap of 100 bp  
\* 100485 102543: contig of 2059 bp in length  
\* 102544 102643: gap of 100 bp  
\* 102644 104695: contig of 2052 bp in length  
\* 104696 104795: gap of 100 bp  
\* 104796 107219: contig of 2424 bp in length  
\* 107220 107319: gap of 100 bp  
\* 107320 109415: contig of 2096 bp in length  
\* 109416 109515: gap of 100 bp  
\* 109516 111921: contig of 2406 bp in length  
\* 111922 112021: gap of 100 bp  
\* 112022 114336: contig of 2315 bp in length  
\* 114337 114436: gap of 100 bp  
\* 114437 131340: contig of 16904 bp in length  
\* 131341 131440: gap of 100 bp  
\* 131441 220770: contig of 89330 bp in length.

FEATURES  
source

1. .220770  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone="RP23-457P12"  
/clone\_lib="RPCI-23"  
1. .5304  
/note="assembly\_fragment:02083  
fragment\_chain:1"  
5405. .50486  
/note="assembly\_fragment:01926  
fragment\_chain:1"  
50587. .62563  
/note="assembly\_fragment:02633  
fragment\_chain:2"  
62664. .77675  
/note="assembly\_fragment:01374  
fragment\_chain:2"  
77776. .79885  
/note="assembly\_fragment:00025"  
79986. .82568  
/note="assembly\_fragment:00210"  
82669. .85961  
/note="assembly\_fragment:00374"  
86062. .88312  
/note="assembly\_fragment:00895"

misc\_feature 88413. .92031  
/note="assembly\_fragment:01131"  
92132. .95041  
/note="assembly\_fragment:01194"  
95142. .98148  
/note="assembly\_fragment:01419"  
98249. .100384  
/note="assembly\_fragment:01433"  
100485. .102543  
/note="assembly\_fragment:01499"  
102644. .104695  
/note="assembly\_fragment:01706"  
104796. .107219  
/note="assembly\_fragment:01768"  
107320. .109415  
/note="assembly\_fragment:01818"  
109516. .111921  
/note="assembly\_fragment:02386"  
112022. .114336  
/note="assembly\_fragment:02608"  
114437. .131340  
/note="assembly\_fragment:00320  
fragment\_chain:3"  
131441. .220770  
/note="assembly\_fragment:01180  
fragment\_chain:3  
clone\_end:SP6  
vector\_side:right"  
BASE COUNT 60179 a 46673 c 47577 g 64342 t 1999 others  
ORIGIN

Query Match 66.8%; Score 27.4; DB 2; Length 220770;  
Best Local Similarity 79.5%; Pred No. 0.69;  
Matches 31; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAGTGTCAAGTGTGAGT 39  
||| ||||| ||||| ||| ||| ||||| |||  
Db 170895 GTGAGAGAGAGATGTGTCGTGTGTGTGTGTGTGT 170857

RESULT 15  
AC123144

LOCUS Rattus norvegicus clone CH230-247A12, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION Rattus norvegicus clone CH230-247A12, \*\*\* SEQUENCING IN PROGRESS  
ACCESSION AC123144 160365 bp DNA linear HTG 13-JUL-2002  
VERSION AC123144.2 GI:21671754  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE  
AUTHORS

1 (bases 1 to 160365)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,  
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,  
Honsi,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

```

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Maehiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,
Oraquyne,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 160365)
Worley,K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 160365)
Worley,K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 2, 2002 this sequence version replaced gi:212339896.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GMGC
Center clone name: CH230-247A12
----- Summary Statistics
Sequencing vector: plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 87649 bases at least Q40
Consensus quality: 91706 bases at least Q30
Consensus quality: 94581 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 70 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1367: contig of 1367 bp in length
* 1368 1467: gap of unknown length
* 1468 2625: contig of 1158 bp in length
* 2626 2725: gap of unknown length
* 2726 3774: contig of 1049 bp in length
* 3775 3874: gap of unknown length
* 3875 4881: contig of 1007 bp in length
* 4882 4981: gap of unknown length
* 4982 6251: contig of 1270 bp in length
* 6252 6351: gap of unknown length
* 6352 7708: contig of 1357 bp in length
*
7709 *
7809 *
9342: contig of 1534 bp in length
9343 *
9442: gap of unknown length
9443 *
10882: contig of 1440 bp in length
10883 *
10982: gap of unknown length
12348: contig of 1366 bp in length
12349 *
12448: gap of unknown length
12449 *
14310: contig of 1862 bp in length
14311 *
14410: gap of unknown length
14411 *
15441: contig of 1031 bp in length
15442 *
15541: gap of unknown length
15542 *
16661: contig of 1120 bp in length
16662 *
16761: gap of unknown length
16762 *
17968: contig of 1107 bp in length
17969 *
17969: gap of unknown length
19159: gap of unknown length
19160 *
20292: contig of 1133 bp in length
20292: gap of unknown length
21970: contig of 1578 bp in length
21971 *
22071: gap of unknown length
22071 *
23357: contig of 1287 bp in length
23358 *
23457: gap of unknown length
23458 *
24901: contig of 1444 bp in length
24902 *
25001: gap of unknown length
25002 *
26788: gap of unknown length
26789 *
28242: contig of 1454 bp in length
28243 *
28342: gap of unknown length
28343 *
30254: contig of 1912 bp in length
30254: gap of unknown length
30354: gap of unknown length
30355 *
31448: contig of 1094 bp in length
31449 *
31548: gap of unknown length
31549 *
33111: contig of 1563 bp in length
33112 *
33211: gap of unknown length
33212 *
34324: contig of 1313 bp in length
34324: gap of unknown length
34624: gap of unknown length
34625 *
35653: contig of 1029 bp in length
35654 *
35753: gap of unknown length
35754 *
37170: contig of 1417 bp in length
37171 *
37270: gap of unknown length
37271 *
39270: gap of unknown length
39270 *
40733: contig of 1364 bp in length
40734 *
40833: gap of unknown length
40834 *
41943: contig of 1110 bp in length
41944 *
42043: gap of unknown length
42044 *
43762: contig of 1719 bp in length
43763 *
43862: gap of unknown length
43863 *
45033: contig of 1171 bp in length
45034 *
45133: gap of unknown length
45134 *
46591: contig of 1458 bp in length
46591: gap of unknown length
46592 *
48528: contig of 1837 bp in length
48529 *
48628: gap of unknown length
48629 *
50052: contig of 1424 bp in length
50053 *
50152: gap of unknown length
50153 *
51567: contig of 1415 bp in length
51568 *
51667: gap of unknown length
51668 *
53892: contig of 2225 bp in length
53892: gap of unknown length
53893 *
56468: contig of 2476 bp in length
56469 *
56568: gap of unknown length
56569 *
59770: contig of 3202 bp in length
59771 *
59870: gap of unknown length
59871 *
60990: contig of 1120 bp in length
60991 *
62896: contig of 1806 bp in length
62897 *
62996: gap of unknown length
62997 *
64476: contig of 1480 bp in length
64477 *
64576: gap of unknown length
64577 *
66186: contig of 1610 bp in length
66187 *
66286: gap of unknown length

```

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

\* 66287 69204: contig of 2918 bp in length  
\* 69204: gap of unknown length  
\* 71613: contig of 2309 bp in length  
\* 71614: gap of unknown length  
\* 71713: contig of 2303 bp in length  
\* 71714: gap of unknown length  
\* 74017 74116: gap of unknown length  
\* 74117 75681: contig of 1565 bp in length  
\* 75682 75781: gap of unknown length  
\* 75782 78005: contig of 2224 bp in length  
\* 78006 78105: gap of unknown length  
\* 78106 80866: contig of 2761 bp in length  
\* 80867 80966: gap of unknown length  
\* 80967 83285: contig of 2319 bp in length  
\* 83286 83385: gap of unknown length  
\* 83386 84752: contig of 1367 bp in length  
\* 84753 84852: gap of unknown length  
\* 84853 87468: contig of 2616 bp in length  
\* 87469 87568: gap of unknown length  
\* 87569 90279: contig of 2711 bp in length  
\* 90280 90379: gap of unknown length  
\* 90380 93894: contig of 3515 bp in length  
\* 93895 93995: gap of unknown length  
\* 93995 96920: contig of 2926 bp in length

Query Match 64.4%; Score 26.4; DB 2; Length 160365;  
Best Local Similarity 78.9%; Pred. No. 1.8;  
Matches 30; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTCGTCYCTAAGTGTGAGTGTGAG 38  
|||||  
Db 47937 GTGTGAGAGAGATGTCGTCGTCGTCGTCGTCGAG 47974

Search completed: February 11, 2003, 05:33:25  
Job time : 257.088 secs





```

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5977 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-425-043-1
Query Match 52.2%; Score 21.4; DB 4; Length 5977;
Best Local Similarity 68.3%; Pred. No. 48;
Matches 28; Conservative 1; Mismatches 12; Indels 0; Gaps 0;
QY 1 AGAAACGAGTGGAGGACGACACCTCAGGCAGCCCGGGAG 41
|| ||| ||||| || | | | | |||||
Db 2795 AGGAGGAGTGGGAAGAGTCGTTTCATGTCGCCAGCGGGGAG 2755

RESULT 6
US-09-024-020B-2/c
; Sequence 2, Application US/09024020B
; Patent No. 6030810
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/425,043
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/024,020
; FILING DATE: 16-FEB-1998
; APPLICATION NUMBER: US 60/039,447
; FILING DATE: 26-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: CLARK, JANET P.
; REGISTRATION NUMBER: 34,799
; REFERENCE/DOCKET NUMBER: R0020B-REG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 852-3097
; TELEFAX: (650) 855-5322
; INFORMATION FOR SEQ ID NO: 2:
; FILING DATE: 16-FEB-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,447
; FILING DATE: 26-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: CLARK, JANET P.
; REGISTRATION NUMBER: 34,799
; REFERENCE/DOCKET NUMBER: R0020B-REG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 852-3097
; TELEFAX: (650) 855-5322
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6007 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-024-020B-2
Query Match 52.2%; Score 21.4; DB 3; Length 6007;
Best Local Similarity 68.3%; Pred. No. 48;
Matches 28; Conservative 1; Mismatches 12; Indels 0; Gaps 0;
QY 1 AGAAACGAGTGGAGGACGACACCTCAGGCAGCCCGGGAG 41
|| ||| ||||| || | | | | |||||
Db 2825 AGGAGGAGTGGGAAGAGTCGTTTCATGTCGCCAGCGGGGAG 2785

RESULT 8
US-09-024-020B-7/c
; Sequence 7, Application US/09024020B
; Patent No. 6030810
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/024,020B
; FILING DATE: 16-FEB-1998
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,447
; FILING DATE: 26-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: CLARK, JANET P.
; REGISTRATION NUMBER: 34,799
; REFERENCE/DOCKET NUMBER: R0020B-REG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 852-3097
; TELEFAX: (650) 855-5322
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6007 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-024-020B-2
Query Match 52.2%; Score 21.4; DB 4; Length 6007;
Best Local Similarity 68.3%; Pred. No. 48;
Matches 28; Conservative 1; Mismatches 12; Indels 0; Gaps 0;
QY 1 AGAAACGAGTGGAGGACGACACCTCAGGCAGCCCGGGAG 41
|| ||| ||||| || | | | | |||||
Db 2825 AGGAGGAGTGGGAAGAGTCGTTTCATGTCGCCAGCGGGGAG 2785
```





```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-024-020B-43

Query Match      52.2%; Score 21.4; DB 3; Length 6586;
Best Local Similarity 68.3%; Pred. No. 48;
Matches 28; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY   1 AGAAGCAGTGGAGGACRACCCTCAGCGCCGCCGGGAG 41
    ||||| ||||||| || | : | | | | | |||||
Db 2972 AGGAAGAGTGGAAGAAGTCGTTCATGTCCAGCGCGGGAG 2932

RESULT 11
US-09-425-043-43/c
; Sequence 43, Application US/09425043
; Patent No. 6335172
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/425,043
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/024,020
; FILING DATE: 16-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: CLARK, JANET P.
; REGISTRATION NUMBER: 34,799
; REFERENCE/DOCKET NUMBER: R0020B-REG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 852-3097
; TELEFAX: (650) 855-5322
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6586 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-425-043-43

Query Match      52.2%; Score 21.4; DB 4; Length 6586;
Best Local Similarity 68.3%; Pred. No. 48;
Matches 28; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY   1 AGAAGCAGTGGAGGACRACCCTCAGCGCCGCCGGGAG 41
    ||||| ||||||| || | : | | | | | |||||
Db 2972 AGGAAGAGTGGAAGAAGTCGTTCATGTCCAGCGCGGGAG 2932

RESULT 12
US-09-024-020B-8/c
; Sequence 8, Application US/09024020B
; Patent No. 6030810
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK
; STREET: 3401 HILLVIEW AVENUE, MS A2-250
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94304-1397
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/024,020B
; FILING DATE: 16-FEB-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,447
; FILING DATE: 26-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: CLARK, JANET P.
; REGISTRATION NUMBER: 34,799
; REFERENCE/DOCKET NUMBER: R0020B-REG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 852-3097
; TELEFAX: (650) 855-5322
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6826 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-024-020B-8

Query Match      52.2%; Score 21.4; DB 3; Length 6826;
Best Local Similarity 68.3%; Pred. No. 48;
Matches 28; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY   1 AGAAAGCAGTGGAGGAGACRACCCTCAGCGCCGCCGGGAG 41
    || ||| ||||||| || | : | | | | | |||||
Db 2768 AGGAAGAGTGGAAGAAGTCGTTCATGTCCAGCGCGGGAG 2728

RESULT 13
US-09-425-043-8/c
; Sequence 8, Application US/09425043
; Patent No. 6335172
; GENERAL INFORMATION:
; APPLICANT: DELGADO, STEPHEN G.
; APPLICANT: DIETRICH, PAUL S.
; APPLICANT: FISH, LINDA M.
; APPLICANT: HERMAN, RONALD C.
; APPLICANT: SANGAMESWARAN, LAKSHMI
; TITLE OF INVENTION: NOVEL CLONED TETRODOTOXIN-SENSITIVE
; TITLE OF INVENTION: SODIUM CHANNEL I-SUBUNIT AND A SPLICE VARIANT THEREOF
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JANET PAULINE CLARK

```

STREET: 3401 HILLVIEW AVENUE, MS A2-250  
CITY: PALO ALTO  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 94304-1397  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/425,043  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/024,020  
FILING DATE: 16-FEB-1998  
APPLICATION NUMBER: US 60/039,447  
FILING DATE: 26-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: CLARK, JANET P.  
REGISTRATION NUMBER: 34,799  
REFERENCE/DOCKET NUMBER: R0020B-REG  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 852-3097  
TELEFAX: (650) 855-5322  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6826 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-425-043-8

Query Match 52.2%; Score 21.4; DB 4; Length 6826;  
Best Local Similarity 68.3%; Pred. No. 48;  
Matches 28; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGCGCCGGGAG 41  
||| ||||| ||| ||| ||||| |||||  
Db 2768 AGGAGGAGTGGAGAGTGTTCATGTGCCAGCGGGGAG 2728

RESULT 14  
US-08-188-582-4  
; Sequence 4, Application US/08188582  
; Patent No. 5534410  
; GENERAL INFORMATION:  
; APPLICANT: Tjian, Robert  
; APPLICANT: Comai, Lucio  
; APPLICANT: Dynlacht, Brian D.  
; APPLICANT: Hoey, Timothy  
; APPLICANT: Ruppert, Siegfried  
; APPLICANT: Tanese, Naoko  
; APPLICANT: Wang, Edith  
; APPLICANT: Weinzierl, Robert O.J.  
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/188,582  
; FILING DATE: 28-JAN-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2359 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 49..2160  
; US-08-188-582-4  
Query Match 51.7%; Score 21.2; DB 1; Length 2359;  
Best Local Similarity 72.2%; Pred. No. 51;  
Matches 26; Conservative 1; Mismatches 9; Indels 0; Gaps 0;  
QY 5 ACAGTGGAGGAGGACRACCCCTCAGCGCCGGGGA 40  
||| ||||| ||| ||| ||||| |||||  
Db 927 ACCAGAGGAGGAGGATGATCGGATCGCCGGGA 962  
RESULT 15  
US-08-646-715-4  
; Sequence 4, Application US/08646715  
; Patent No. 5637686  
; GENERAL INFORMATION:  
; APPLICANT: Tjian, Robert  
; APPLICANT: Comai, Lucio  
; APPLICANT: Dynlacht, Brian D.  
; APPLICANT: Hoey, Timothy  
; APPLICANT: Ruppert, Siegfried  
; APPLICANT: Tanese, Naoko  
; APPLICANT: Wang, Edith  
; APPLICANT: Weinzierl, Robert O.J.  
; TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/646,715  
; FILING DATE: 09-MAY-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/188,582  
; FILING DATE: 28-JAN-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Osman, Richard A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989

```
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2359 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 49..2160
US-08-646-715-4

Query Match      51.7%; Score 21.2; DB 1; Length 2359;
Best Local Similarity 72.2%; Pred. No. 51;
Matches 26; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY  5  AGCAGTGGAGGAGGACRACCTCAGGCAGCCCGGGA 40
    | | | | | | | | | | | | | | | | | | | |
Db  927 ACCAGAGAGGAGGAGGATGATCCGGATGCCCGGA 962
```

Search completed: February 11, 2003, 05:59:51  
Job time : 1049.12 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 03:31:30 ; Search time 3.72137 Seconds  
(without alignments)  
5222.300 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_600\_640

Perfect score: 41

Sequence: 1 agaagcagtgaggagac.....accctcagcgagccgggag 41

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries.

Database : Published Applications\_NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*

10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*

12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*

13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*

14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24.6	60.0	159	10	US-09-864-761-25857
2	24.6	60.0	555	10	US-09-864-761-9308
3	24.6	60.0	2636	10	US-09-789-404-1
4	22.4	54.6	3314	10	US-09-764-864-490
5	22	53.7	1845	10	US-09-778-844-64
6	21.8	53.2	195	10	US-09-864-761-18173
7	21.8	53.2	195	10	US-09-864-761-19007
8	21.8	53.2	432	10	US-09-876-889-223
9	21.8	53.2	461	10	US-09-864-761-2269
10	21.8	53.2	462	10	US-09-864-761-2795
11	21.8	53.2	620	10	US-09-925-297-303
12	21.8	53.2	906	10	US-09-925-297-302
13	21.8	53.2	1006	9	US-09-964-899-22
14	21.8	53.2	17252	10	US-09-764-860-1102
15	21.8	53.2	76798	10	US-09-880-107-3949
16	21.4	52.2	267	10	US-09-923-876-460
17	21.4	52.2	170834	10	US-09-835-232-7
18	21.2	51.7	1143	10	US-09-822-849A-95
19	21	51.2	183	10	US-09-864-761-26896

ALIGNMENTS

RESULT 1

US-09-864-761-25857  
; Sequence 25857, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO  
; FILE REFERENCE: Aecmicsa-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30

Sequence 1948, Ap  
Sequence 10261, A  
Sequence 1652, Ap  
Sequence 1652, Ap  
Sequence 1, Appli  
Sequence 11, Appli  
Sequence 18, Appli  
Sequence 2188, Ap  
Sequence 1986, Ap  
Sequence 2036, Ap  
Sequence 1986, Ap  
Sequence 2036, Ap  
Sequence 1139, Ap  
Sequence 5935, Ap  
Sequence 2195, Ap  
Sequence 9, Appli  
Sequence 5, Appli  
Sequence 1, Appli  
Sequence 16, Appli  
Sequence 1713, Ap  
Sequence 18234, A  
Sequence 87, Appli  
Sequence 1475, Ap  
Sequence 1364, Ap  
Sequence 839, App  
Sequence 3565, Ap

```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25857
; LENGTH: 159
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL035681.13
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.3
; OTHER INFORMATION: NT HIT: 002972.1, EVALUE 9.20e-02
; OTHER INFORMATION: EST_HUMAN HIT: BF448000.1, EVALUE 2.00e-84
; OTHER INFORMATION: SWISSPROT HIT: P16356, EVALUE 2.00e-01
US-09-864-761-25857
```

```
Query Match 60.0%; Score 24.6; DB 10; Length 159;
Best Local Similarity 73.2%; Pred. No. 2.9;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 AGAAGCAGTGGAGGAGGACRACCTCAGCGAGCGCGGAG 41
||||| 1 || |||||:||||||| ||| || ||
Db 49 AGAAGCCCTGCTGGAGGAGGACCCCTCAGGTGCCAGGAG 89
```

## RESULT 2

```
US-09-864-761-9308
; Sequence 9308, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9308
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL035681.13
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
US-09-864-761-9308
```

```
Query Match 60.0%; Score 24.6; DB 10; Length 555;
Best Local Similarity 73.2%; Pred. No. 3.3;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 1 AGAAGCAGTGGAGGAGGACRACCTCAGCGAGCGCGGAG 41
||||| 1 || |||||:||||||| ||| || ||
Db 298 AGAAGCCCTGCTGGAGGAGGACCCCTCAGGTGCCAGGAG 338
```

## RESULT 3

```
US-09-789-404-1/c
; Sequence 1, Application US/09789404
; Patent No. US20020025554A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran
; TITLE OF INVENTION: NOVEL LEUCINE RICH REPEAT-CONTAINING MOLECULES AND USES THEREF
; FILE REFERENCE: 10448/008001
; CURRENT APPLICATION NUMBER: US/09/789,404
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/456,592
; PRIOR FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2636
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (33)...(2414)
; NAME/KEY: misc.feature
; LOCATION: (1)...(2636)
; OTHER INFORMATION: n = A,T,C or G
US-09-789-404-1

Query Match 60.0%; Score 24.6; DB 10; Length 2636;
Best Local Similarity 73.2%; Pred. No. 3.8;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGCGAGCGCGGAG 41
```

```
|||| | || |||||:||||||| ||| || ||
Db 2533 AGAAGCCCTGCTGGAGGACGACCCTCAGGGTCCAGGAAG 2493

RESULT 4
US-09-764-864-490
; Sequence 490, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 490
; LENGTH: 3314
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (116)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (244)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (293)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (305)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (403)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (439)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (448)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3305)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (3311)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-864-490

Query Match 54.6%; Score 22.4; DB 10; Length 3314;
Best Local Similarity 76.5%; Pred. No. 23;
Matches 26; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 AAAGCAGTGGAGGAGGACRACCTCAGGCAGCCC 36
||| || |||||:|||| ||||| |||||
Db 1888 ATGCCCGAGGAGGAGGAACCTTCAGGCAGCCC 1921

RESULT 5
US-09-778-844-64
; Sequence 64, Application US/09778844
; Patent No. US20020150971A1
; GENERAL INFORMATION:
; APPLICANT: JOHANSEN, JEANETTE ELISABETH
; APPLICANT: SCHALLING, MARTIN
; TITLE OF INVENTION: NUCLEIC ACIDS AND POLYPEPTIDES FOR CONTROLLING FOOD
; FILE REFERENCE: 030307/0195
```

```
; CURRENT APPLICATION NUMBER: US/09/778,844
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 1845
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Chp, EMBL No. US20020150971A1
; OTHER INFORMATION: q9zly0
US-09-778-844-64

Query Match 53.7%; Score 22; DB 10; Length 1845;
Best Local Similarity 70.0%; Pred. No. 31;
Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2 GAAAGCAGTGGAGGAGGACRACCTCAGGCAGCCCCGGGAG 41
|| ||| ||||| ||| ||| | ||| |||| |||
Db 1504 GACAGCTGTGGAGAGGCCGACACCTGGGCTGCCCTCGAG 1543

RESULT 6
US-09-864-761-18173
; Sequence 18173, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
```

```
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18173
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO 284718.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: NT HIT: AF240786.1, EVALUE 1.00e-106
; OTHER INFORMATION: EST_HUMAN HIT: A1281103.1, EVALUE 1.00e-33
; OTHER INFORMATION: SWISSPROT HIT: P30711, EVALUE 5.00e-33
US-09-864-761-18173
```

```
Query Match 53.2%; Score 21.8; DB 10; Length 195;
Best Local Similarity 70.7%; Pred. No. 29;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
```

```
QY 1 AGAAGCAGTGGAGGAGGACRACCTCAGGCAGCCCGGGAG 41
|| ||||| ||||| |||| |||| |||| |||| |||| ||||
Db 68 AGGCAGCAGTGGGGAGGACCTCTTCCAGGAGGCCCATGAG 108
```

```
RESULT 7
US-09-864-761-19007/c
; Sequence 19007, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19007
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000351.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 12
; OTHER INFORMATION: SWISSPROT HIT: P30711, EVALUE 5.00e-33
; OTHER INFORMATION: EST_HUMAN HIT: A1281103.1, EVALUE 1.00e-106
; OTHER INFORMATION: NT HIT: gi4504184, EVALUE 1.00e-106
US-09-864-761-19007
```

```
Query Match 53.2%; Score 21.8; DB 10; Length 195;
Best Local Similarity 70.7%; Pred. No. 29;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
```

```
QY 1 AGAAGCAGTGGAGGAGGACRACCTCAGGCAGCCCGGGAG 41
|| ||||| ||||| |||| |||| |||| |||| |||| ||||
Db 128 AGGCAGCAGTGGGGAGGACCTCTTCCAGGAGGCCCATGAG 88
```

```
RESULT 8
US-09-876-889-223
; Sequence 223, Application US/09876889
; Patent No. US20020076715A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE REFERENCE: 210121.466C3
; CURRENT APPLICATION NUMBER: US/09/876,889
; CURRENT FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 353
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 223
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-876-889-223
```

```
Query Match 53.2%; Score 21.8; DB 10; Length 432;
Best Local Similarity 70.7%; Pred. No. 31;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
```

```
QY 1 AGAAGCAGTGGAGGAGGACRACCTCAGGCAGCCCGGGAG 41
|| ||||| ||||| |||| |||| |||| |||| |||| ||||
Db 332 AGGCAGCAGTGGGGAGGACCTCTTCCAGGAGGCCCATGAG 372
```

```
RESULT 9
US-09-864-761-2269/c
```



```
; Sequence 2269, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2269
; LENGTH: 461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AP000351.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; US-09-864-761-2269

Query Match 53.2%; Score 21.8; DB 10; Length 461;
Best Local Similarity 70.7%; Pred. No. 32;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGAAGCAGCTGGAGGAGACACCTCTCCAGGAGGCCCATGAG 41
```

```
Db 374 AGCAGCAGTGGGGAGGAGACCTCTTCAGGAGGCCCATGAG 334
```

```
RESULT 10
US-09-864-761-2795
; Sequence 2795, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2795
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO 284718.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; US-09-864-761-2795
```

```
Query Match          53.2%; Score 21.8; DB 10; Length 462;
Best Local Similarity 70.7%; Pred. No. 32;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGAAAGCAGTGGAGGAGGACRACCCCTCAGGCGAGCCCGGGAG 41
   || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 366 AGGCAGCAGTGGGGAGGACCTCTTCCAGGAGGCCCATGAG 406

RESULT 11
US-09-925-297-303
; Sequence 303, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 303
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (125)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (620)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-303

Query Match          53.2%; Score 21.8; DB 10; Length 620;
Best Local Similarity 70.7%; Pred. No. 32;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGAAAGCAGTGGAGGAGGACRACCCCTCAGGCGAGCCCGGGAG 41
   || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 133 AGGCAGCAGTGGGGAGGACCTCTTCCAGGAGGCCCATGAG 173

RESULT 12
US-09-925-297-302
; Sequence 302, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 302
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (584)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
```

```
; LOCATION: (627)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (863)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-297-302

Query Match          53.2%; Score 21.8; DB 10; Length 906;
Best Local Similarity 70.7%; Pred. No. 34;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGAAAGCAGTGGAGGAGGACRACCCCTCAGGCGAGCCCGGGAG 41
   || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 639 AGGCAGCAGTGGGGAGGACCTCTTCCAGGAGGCCCATGAG 679

RESULT 13
US-09-964-899-22
; Sequence 22, Application US/09964899
; Patent No. US2002017446A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Dalia et al.
; TITLE OF INVENTION: Identification of Genes Involved in
; FILE REFERENCE: 4-31612 A
; CURRENT APPLICATION NUMBER: US/09/964,899
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/236,893
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/298,309
; PRIOR FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 1006
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-964-899-22

Query Match          53.2%; Score 21.8; DB 9; Length 1006;
Best Local Similarity 70.7%; Pred. No. 34;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 1 AGAAAGCAGTGGAGGAGGACRACCCCTCAGGCGAGCCCGGGAG 41
   || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 596 AGGCAGCAGTGGGGAGGACCTCTTCCAGGAGGCCCATGAG 636

RESULT 14
US-09-764-860-1102/c
; Sequence 1102, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1102
; LENGTH: 17252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1102

Query Match          53.2%; Score 21.8; DB 10; Length 17252;
Best Local Similarity 78.8%; Pred. No. 45;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 AGAAAGCAGTGGAGGAGGACRACCCCTCAGGCGAG 33
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```





GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:43:59 ; Search time 16.6596 Seconds  
(without alignments)  
5542.256 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_860\_900

Perfect score: 41

Sequence: 1 ggtgagagagaatgtgtgc.....ctaagtgcagtgtgagtct 41

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.6	99.0	1669	AAH26169	Human cytochrome P
2	40.6	99.0	1669	AAH26179	Human cytochrome P
3	40.6	99.0	1680	AAD34214	Human CYP2D6 gene
4	40.6	99.0	6472	ABQ72215	Human CYP2D6 gene,
5	40.6	99.0	6472	ABQ72364	Human CYP2D6 gene,
6	40.6	99.0	9432	AAD34213	Human cytochrome p
7	32.6	79.5	5884	ABK39958	Human chemically p
8	32.6	79.5	5884	ABL32556	Human immune syste
9	25	61.0	20015	ABA20740	Human nervous syst

C	10	23.2	56.6	280	24	ABN22337	Human ORFX polynuc
	11	23	56.1	248	21	AAC22210	Human secreted pro
	12	22.6	55.1	75384	22	AAK85590	Human immune/halema
	13	22	53.7	468	21	AAA43852	Human secreted exp
	14	22	53.7	479	18	AAT91307	Human H83-22 secre
	15	22	53.7	479	18	AAV00431	5' fragment of clo
	16	22	53.7	570	23	ABV57148	Human prostate exp
	17	22	53.7	1040	24	ABL64398	Stomach cancer rel
	18	22	53.7	1078	18	AAT91308	Human H83-22 secre
	19	22	53.7	2013	22	AAH26535	Human hPEPT1 gene
	20	22	53.7	7599	22	AAI04162	Human reproductive
	21	22	53.7	14985	17	AAT43574	Human box-dependen
	22	22	53.7	14985	19	AAV15693	Human Bin1 gene.
	23	22	53.7	23677	20	AAI16323	Partial human geno
	24	21.8	53.2	485	22	ABA43852	Human breast cell
	25	21.8	53.2	485	22	ABA54314	Human foetal liver
	26	21.8	53.2	485	22	ABA24072	Probe #2538 for ge
	27	21.8	53.2	485	22	AAK02596	Human brain expres
	28	21.8	53.2	485	22	AAK28030	Human bone marrow
	29	21.8	53.2	485	22	AAI12612	Probe #2650 used t
	30	21.8	53.2	485	22	AAI33964	Probe #2545 for ge
	31	21.8	53.2	485	22	AAI02515	Probe #2506 used t
	32	21.8	53.2	485	24	ABS02512	Human genome-deriv
	33	21.8	53.2	576	23	ABV60342	Human prostate exp
	34	21.4	52.2	349	23	ABV00080	Human prostate exp
	35	21.4	52.2	349	23	ABV09249	Human prostate exp
	36	21.4	52.2	397	23	ABV30425	Human prostate exp
	37	21.4	52.2	397	23	ABV39400	Human prostate exp
	38	21.4	52.2	895	22	AA542055	Genomic sequence #
	39	21.4	52.2	895	22	AAI02801	Human reproductive
	40	21.4	52.2	2430	24	ABL90159	Human polynucleoti
	41	21.4	52.2	30032	22	ABAI7086	Human nervous syst
	42	21.2	51.7	956	21	AAC50116	Arabidopsis thalia
	43	21.2	51.7	960	21	AAC33700	Arabidopsis thalia
	44	21.2	51.7	4474	23	ABL26934	Drosophila melanog
	45	21	51.2	339	24	ABL85097	Human ovarian canc

ALIGNMENTS

RESULT 1  
AAH26169  
ID AAH26169 standard; DNA; 1669 BP.  
XX  
AC AAH26169;  
XX  
DT 17-SEP-2001 (first entry)  
XX  
DE Human cytochrome P450 CYP2D6 gene promoter region.  
XX  
KW Cytochrome P450; CYP2D6; promoter; drug metabolism; human;  
KW diagnosis; therapy; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT primer\_bind  
FT complement (14...36)  
FT /tag= a  
FT /note= amplification primer upf14  
FT primer\_bind  
FT 337...358  
FT /tag= b  
FT /note= sequence primer R1  
FT primer\_bind  
FT 493...514  
FT /tag= c  
FT /note= sequencing primer R2  
FT primer\_bind  
FT complement (565...577)  
FT /tag= d  
FT /note= sequencing primer F2  
FT primer\_bind  
FT 602...620  
FT /tag= e  
FT /note= sequencing primer R3  
FT complement (968...988)

FT /\*tag= f  
FT /note= "sequencing primer F3"  
FT 1124..1143  
FT /\*tag= g  
FT /note= "sequencing primer R4"  
FT 1605..1623  
FT /\*tag= h  
FT /note= "sequencing primer R5"  
FT 1650..1669  
FT /\*tag= i  
FT /note= "amplification primer upr1669"  
FT 1532..1619  
FT /\*tag= j  
FT 1620..1669  
FT /\*tag= k  
FT /partial  
FT /note= "5' region of CYP2D6 coding region"  
XX

PN WO200155432-A2.

XX 02-AUG-2001.

XX 30-JAN-2001; 2001WO-EP00954.

XX 31-JAN-2000; 2000EP-0101889.

XX (EPID-) EPIDAURUS BIOTECHNOLOGIE AG.

XX Raimundo S, Zanger U;

XX WPI; 2001-457734/49.

XX A polynucleotide capable of hybridizing to CYP2D6 promoter useful for  
PT the optimization of drug therapies using substrates of cytochrome P-450  
PT -

XX Claim 1; Fig 1; 41pp; English.

XX The present sequence is that of the promoter region of the human  
CC cytochrome P450 CYP2D6 gene. The promoter region was amplified  
CC by PCR from leucocyte DNA of over 50 individuals, and sequenced.  
CC 8 previously unknown single nucleotide polymorphisms (SNP) were  
CC identified. These were at: base 36 (base -1584 according to the  
CC Human Cytochrome P450 Allele Nomenclature), where the SNP was C to  
CC G, occurring at an estimated frequency of approximately 20% in the  
CC whole population, and resulting in increased enzyme activity;  
CC position 194 (-1426), C to T, approximately 20% frequency; neutral  
CC function; position 365 (-1235), A to G, approximately 50% frequency,  
CC neutral function; position 620 (-1000), G to A, approximately 20%  
CC frequency; neutral function; position 880 (-740), C to T,  
CC approximately 30% frequency; unknown function; position 940 (-680),  
CC G to A, approximately 30% frequency; unknown function; 1255 (-365),  
CC G to A, rare, unknown function; and 1298 (-322), T to C, rare,  
CC unknown function. The C to G mutation at -1584 bp is strongly  
CC associated with lower metabolic ratios, and a molecular variant  
CC polynucleotide having G at this position is claimed (see AAH26179).  
CC The invention provides a method of diagnosing a disorder related to  
CC reduced or enhanced capacity for clearance of CYP2D6 substrates  
CC (antiarrhythmic, beta-adrenergic receptor-antagonist, tricyclic  
CC antidepressant, selective serotonin reuptake inhibitor, neuroleptic,  
CC opiate, cytotatic or amphetamine), or susceptibility to such a  
CC disorder. By determining the presence of a mutation in the CYP2D6  
CC promoter. The strong association of the common C to G mutation at  
CC -1584 bp with increased enzyme activity significantly improves the  
CC correlation between genotype and phenotype in the CYP2D6 polymorphism.  
CC Testing for the mutation will allow the identification of intermediate  
CC metabolizers and therefore allow quantitative predictions to be made  
CC on in vivo drug metabolism capacity, thus providing a very potent  
CC tool for improving the therapy of diseases with drugs that are  
CC targets of the CYP2D6 gene product.

XX Sequence 1669 BP; 413 A; 376 C; 534 G; 338 T; 8 other;

Query Match 99.0%; Score 40.6; DB 22; Length 1669;  
Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGAAATGTGTCYCTAAGTGTGTCAGTGTGAGTCT 41  
|||||  
Db 860 GTGTGAGAGAGAAATGTGTCYCTAAGTGTGTCAGTGTGAGTCT 900  
|||||

RESULT 2

AAH26179

ID AAH26179 standard; DNA; 1669 BP.

XX AC AAH26179;

XX DT 17-SEP-2001 (first entry)

XX DE Human cytochrome P450 CYP2D6 gene promoter (G mutation at -1584 bp).

XX KW Cytochrome P450; CYP2D6; promoter; drug metabolism; human;  
XX KW diagnosis; therapy; single nucleotide polymorphism; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
FT variation replace(36,G)  
FT /\*tag= a

FT FT 5'UTR /frequency= "20%"

FT FT 1532..1619

FT FT /\*tag= b

FT CDS 1620..1669

FT FT /\*tag= c

FT FT /partial

FT FT /note= "5' region of CYP2D6 coding region"

XX

PN WO200155432-A2.

XX 02-AUG-2001.

XX 30-JAN-2001; 2001WO-EP00954.

XX 31-JAN-2000; 2000EP-0101889.

XX (EPID-) EPIDAURUS BIOTECHNOLOGIE AG.

XX Raimundo S, Zanger U;

XX WPI; 2001-457734/49.

XX A polynucleotide capable of hybridizing to CYP2D6 promoter useful for  
PT the optimization of drug therapies using substrates of cytochrome P-450  
PT -

XX Claim 1(a); Page -; 41pp; English.

XX The present sequence is that of the promoter region of the human  
CC cytochrome P450 CYP2D6 gene, which includes G at position 36  
CC (base -1584 according to the Human Cytochrome P450 Allele  
CC Nomenclature). The presence of C at position -1584 bp is a marker  
CC for low enzyme activity, whereas there is strong association of G  
CC at position -1584 bp with increased enzyme activity. The C to G  
CC single nucleotide polymorphism occurs in approximately 20% of the  
CC population. The invention provides a method of diagnosing a  
CC disorder related to reduced or enhanced capacity for clearance of  
CC CYP2D6 substrates (antiarrhythmic, beta-adrenergic receptor  
CC antagonist, tricyclic antidepressant, selective serotonin reuptake  
CC inhibitor, neuroleptic, opiate, cytotatic or amphetamine), or  
CC susceptibility to such a disorder, by determining the presence of  
CC a mutation in the CYP2D6 promoter. The novel variant forms of the  
CC CYP2D6 gene provided by the invention provide the potential for the  
CC development of a pharmacodynamic profile of drugs for a given  
CC patient. The finding and characterization of variations in the  
CC CYP2D6 gene, and diagnostic tests for the discrimination of

CC different alleles in human individuals, provide a very potent tool  
CC for improving the therapy of diseases with drugs that are targets  
CC of the CYP2D6 gene production, and whose metabolism is therefore  
CC dependent on CYP2D6 activity.  
CC Note: The present sequence is not shown in the specification but is  
CC derived from the CYP2D6 promoter sequence given in the Sequence  
CC Listing (see AAH26169).  
XX  
SQ Sequence 1669 BP; 413 A; 376 C; 535 G; 338 T; 7 other;  
Query Match 99.0%; Score 40.6; DB 22; Length 1669;  
Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTGTGAGAGAGAAATGTCGTCYCTAAGTGTCAAGTGTGAGTCT 41  
|||||  
Db 860 GTGTGAGAGAGAAATGTCGTCYCTAAGTGTCAAGTGTGAGTCT 900  
RESULT 3  
AAD34214  
ID AAD34214 standard; DNA: 1680 BP.  
XX  
AC AAD34214;  
XX  
DT 16-JUL-2002 (first entry)  
XX  
DE Human CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX  
KW Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; polymorphism; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 36  
FT /\*tag= a  
FT /note= "Polymorphic site"  
FT 194  
FT /\*tag= b  
FT /note= "Polymorphic site"  
FT 385  
FT /\*tag= c  
FT /note= "Polymorphic site"  
FT 620  
FT /\*tag= d  
FT /note= "Polymorphic site"  
FT 880  
FT /\*tag= e  
FT /note= "Polymorphic site"  
FT 942  
FT /\*tag= f  
FT /note= "Polymorphic site"  
FT 1255  
FT /\*tag= g  
FT /note= "Polymorphic site"  
XX  
PN WO200218638-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 27-AUG-2001; 2001WO-IB01544.  
XX  
PR 30-AUG-2000; 2000GB-0021286.  
XX  
PA (GEMI-) GEMINI GENOMICS PLC.  
XX  
PI Risinger C, Andersson MK, Lewander T, Oliasson E;  
XX  
DR WPI; 2002-329785/36.  
XX  
PT New sequence determination oligonucleotides, useful for detecting

PT polymorphic sites in a 5' flanking region of a CYP2D6 gene, as  
PT hybridization probes, as components of diagnostic assays, or in  
PT ligase-based sequence determination -  
XX  
PS Claim 1; Fig 2; 63pp; English.  
XX  
CC The invention relates to sequence determination oligonucleotides for  
CC detecting polymorphic sites in a 5' flanking region of cytochrome P450  
CC 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many  
CC different xenobiotics. Human CYP2D6 gene is located on chromosome 22.  
CC The oligonucleotides may be used as in situ hybridisation probes, in  
CC ligase-based sequence determination, as components of diagnostic assays,  
CC as probes in sequence determination methods based on mismatches, as  
CC hybridisation-based diagnostic assays, and as components of diagnostic  
CC microarray. CYP2D6 is useful to predict variations in an individual's  
CC ability to metabolise certain drugs. The present sequence is human  
CC CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX  
SQ Sequence 1680 BP; 413 A; 379 C; 539 G; 342 T; 7 other;  
Query Match 99.0%; Score 40.6; DB 24; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 5.2e-07;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GTGTGAGAGAGAAATGTCGTCYCTAAGTGTCAAGTGTGAGTCT 41  
|||||  
Db 860 GTGTGAGAGAGAAATGTCGTCYCTAAGTGTCAAGTGTGAGTCT 900  
RESULT 4  
ABQ72215  
ID ABQ72215 standard; DNA: 6472 BP.  
XX  
AC ABQ72215;  
XX  
DT 02-SEP-2002 (first entry)  
XX  
DE Human CYP2D6 gene, SEQ ID NO:1 version #1.  
XX  
KW Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;  
KW chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;  
KW antiarrhythmic; arrhythmia; adrenoreceptor antagonist; hypertension;  
KW tricyclic antidepressant; procainamide; drug induced lupus syndrome;  
KW environmentally linked disease; Parkinson's disease; haplotyping;  
KW genotyping; haplotype; genetic variant; single nucleotide polymorphism;  
KW SNP; drug screening; drug discovery; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT variation replace (636, A)  
FT /\*tag= a  
FT /label= PS1  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT variation replace (678, C)  
FT /\*tag= b  
FT /label= PS2  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT variation replace (769, C)  
FT /\*tag= c  
FT /label= PS3  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT variation replace (776, G)  
FT /\*tag= d  
FT /label= PS4  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT variation replace (825, A)  
FT /\*tag= e  
FT /label= PS5

```

FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT replace (915, C)
FT FT /*tag= f
FT FT /label= PS6
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT 1001..5217
FT FT /*tag= g
FT FT /product= "CYP2D6"
FT FT 1001..1180
FT FT /*tag= h
FT FT /number= 1
FT FT replace (1019, A)
FT FT /*tag= i
FT FT /label= PS7
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT given as R in the specification; causes the
FT FT amino acid substitution V7M"
FT FT replace (1031, A)
FT FT /*tag= j
FT FT /label= PS8
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT given as R in the specification; causes the
FT FT amino acid substitution V11M"
FT FT replace (1100, T)
FT FT /*tag= k
FT FT /label= PS9
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT given as Y in the specification; causes the
FT FT amino acid substitution P34S"
FT FT 1181..1183
FT FT /*tag= l
FT FT /number= 1
FT FT replace (1827, C)
FT FT /*tag= m
FT FT /label= PS10
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as S in the specification"
FT FT replace (1843, G)
FT FT /*tag= n
FT FT /label= PS11
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as K in the specification"
FT FT 1884..2055
FT FT /*tag= o
FT FT /number= 2
FT FT replace (1966, A)
FT FT /*tag= p
FT FT /label= PS12
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as R in the specification; causes the
FT FT amino acid substitution R88H"
FT FT replace (1974, A)
FT FT /*tag= q
FT FT /label= PS13
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT given as M in the specification; causes the
FT FT amino acid substitution L91M"
FT FT replace (1984, G)
FT FT /*tag= r
FT FT /label= PS14
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as R in the specification; causes the
FT FT amino acid substitution H94R"
FT FT replace (1997, G)
FT FT /*tag= s
FT FT /label= PS15
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as S in the specification"
FT FT replace (2014, C)
FT FT /*tag= t
FT FT /label= PS16
FT FT

FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as Y in the specification; causes the
FT FT amino acid substitution V104A"
FT FT replace (2022, T)
FT FT /*tag= u
FT FT /label= PS17
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as W in the specification; together
FT FT with PS18 causes the amino acid substitution
FT FT T107F"
FT FT 2056..2605
FT FT /*tag= v
FT FT /label= PS18
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as Y in the specification; together
FT FT with PS17 causes the amino acid substitution
FT FT T107F"
FT FT replace (2028, G)
FT FT /*tag= w
FT FT /label= PS19
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as R in the specification; causes the
FT FT amino acid substitution I109V"
FT FT replace (2036, C)
FT FT /*tag= x
FT FT /label= PS20
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as Y in the specification"
FT FT replace (2039, T)
FT FT /*tag= y
FT FT /label= PS21
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT given as Y in the specification"
FT FT 2056..2605
FT FT /*tag= z
FT FT /number= 2
FT FT /cons_splice= (5'site:NO, 3'site:YES)
FT FT replace (2062, G)
FT FT /*tag= aa
FT FT /label= PS22
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as R in the specification"
FT FT replace (2067, G)
FT FT /*tag= ab
FT FT /label= PS23
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as K in the specification"
FT FT replace (2118, T)
FT FT /*tag= ac
FT FT /label= PS24
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as Y in the specification"
FT FT replace (2170, A)
FT FT /*tag= ad
FT FT /label= PS25
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT given as R in the specification"
FT FT replace (2179, C)
FT FT /*tag= ae
FT FT /label= PS26
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as S in the specification"
FT FT 2606..2758
FT FT /*tag= af
FT FT /number= 3
FT FT replace (2611, A)
FT FT /*tag= ag
FT FT /label= PS27
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT given as W in the specification; causes the
FT FT amino acid substitution F120I"
FT FT replace (2635, C)
FT FT /*tag= ah
FT FT

```



```
FT /label= PS28
FT /note= "Novel single nucleotide polymorphism (SNP);
FT given as Y in the specification; causes the
FT amino acid substitution W128R"
FT variation
FT replace (2659, A)
FT /tag= ai
FT /label= PS29
FT /note= "Novel single nucleotide polymorphism (SNP);
FT given as R in the specification; together
FT with PS30 causes the amino acid substitution
FT V136I"
FT variation
FT replace (2661, C)

Query Match 99.0%; Score 40.6; DB 24; Length 6472;
Best Local Similarity 97.6%; Pred. No. 7.1e-07;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTCGTCYCTAAGTCGAGTGTGAGTCT 41
Db 241 GTGTGAGAGAGATGTCGCCCTAAGTCGAGTGTGAGTCT 281

RESULT 5
ID ABQ72364
XX AC ABQ72364 standard; DNA; 6472 BP.
XX AC ABQ72364;
XX 02-SEP-2002 (first entry)
XX Human CYP2D6 gene, SEQ ID NO:1 version #2.
XX Human: cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;
KW chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;
KW antirhythmic; arrhythmia; adrenoceptor antagonist; hypertension;
KW triacyclic antidepressant; procainamide; drug induced lupus syndrome;
KW environmentally linked disease; Parkinson's disease; haplotyping;
KW genotyping; haplotype; genetic variant; single nucleotide polymorphism;
KW SNP; drug screening; drug discovery; gene; ds.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH variation replace (636, A)
FT /tag= a
FT /label= PS1
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (678, C)
FT /tag= b
FT /label= PS2
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (769, C)
FT /tag= c
FT /label= PS3
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (776, G)
FT /tag= d
FT /label= PS4
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (825, A)
FT /tag= e
FT /label= PS5
FT /note= "Known single nucleotide polymorphism (SNP)"
FT variation replace (915, C)
FT /tag= f
FT /label= PS6
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT CDS 1001..5217
FT /tag= g
FT /product= "CYP2D6"
FT exon 1001..1180
FT /tag= h
FT /number= 1
```

```
FT variation
FT /tag= i
FT /label= PS7
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V7M"
FT variation
FT replace (1031, A)
FT /tag= j
FT /label= PS8
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V11M"
FT variation
FT replace (1100, T)
FT /tag= k
FT /label= PS9
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution P34S"
FT intron 1181..1883
FT /tag= l
FT /number= 1
FT variation replace (1827, C)
FT /tag= m
FT /label= PS10
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (1843, G)
FT /tag= n
FT /label= PS11
FT /note= "Known single nucleotide polymorphism (SNP)"
FT exon 1884..2055
FT /tag= o
FT /number= 2
FT variation replace (1966, A)
FT /tag= p
FT /label= PS12
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution R88H"
FT variation replace (1974, A)
FT /tag= q
FT /label= PS13
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution L91M"
FT variation replace (1984, G)
FT /tag= r
FT /label= PS14
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution H94R"
FT variation replace (1997, G)
FT /tag= s
FT /label= PS15
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (2014, C)
FT /tag= t
FT /label= PS16
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V104A"
FT variation replace (2022, T)
FT /tag= u
FT /label= PS17
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS17 causes the amino acid
FT substitution T107F"
FT variation replace (2023, T)
FT /tag= v
FT /label= PS18
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS17 causes the amino acid
FT substitution T107F"
FT variation replace (2028, G)
FT /tag= w
FT /label= PS19
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution I109V"
FT variation replace (2036, C)
FT /tag= x
FT /label= PS20
```

```
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace (2039, T)
FT /*tag= Y
FT /label= PS21
FT /note= "Known single nucleotide polymorphism (SNP)"
FT 2056..2605
FT /*tag= z
FT /number= 2
FT /cons.splice= (5'site:NO, 3'site:YES)
FT replace (2062, G)
FT /*tag= aa
FT /label= PS22
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace (2067, G)
FT /*tag= ab
FT /label= PS23
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace (2118, T)
FT /*tag= ac
FT /label= PS24
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT replace (2170, A)
FT /*tag= ad
FT /label= PS25
FT /note= "Known single nucleotide polymorphism (SNP)"
FT replace (2179, C)
FT /*tag= ae
FT /label= PS26
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT 2606..2758
FT /*tag= af
FT /number= 3
FT replace (2611, A)
FT /*tag= ag
FT /label= PS27
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution F120I"
FT replace (2635, C)
FT /*tag= ah
FT /label= PS28
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution W128R"
FT replace (2659, A)
FT /*tag= ai
FT /label= PS29
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS30 causes the amino acid
FT substitution V136I"
FT replace (2661, C)
FT /*tag= aj
FT /label= PS30
FT /note= "Known single nucleotide polymorphism (SNP);
FT together with PS29 causes the amino acid
FT substitution V136I"
FT replace (2704, G)
FT /*tag= ak
FT /label= PS31
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution Q151E"
FT replace (2716, A)
FT /*tag= al
FT /label= PS32
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution E155R"
FT 2759..2846
FT /*tag= am
FT /number= 3
FT replace (2846, A)
FT /*tag= an
FT /label= PS33
FT /note= "Known single nucleotide polymorphism (SNP)"
FT 2847..3007
FT /*tag= ao
```

```
FT intron
FT /number= 4
FT 3008..3440
FT /*tag= ap
FT /number= 4
FT replace (3292, A)

Query Match          99.0%; Score 40.6; DB 24; Length 6472;
Best Local Similarity 97.6%; Pred. No. 7.1e-07;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGAAATGTGTCYCTAAGTGTGAGTGTGAGTCT 41
   ||||||||||||||||||||||||||||||||||||||||
Db 241 GTGTGAGAGAGAAATGTGTCCTAAGTGTGAGTGTGAGTCT 281

RESULT 6
AAD34213
ID AAD34213 standard; DNA; 9432 BP.
XX
AC AAD34213;
XX
DT 16-JUL-2002 (first entry)
XX
DE Human cytochrome P450 2D6 (CYP2D6) gene.
XX
KW Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;
KW ligase-based sequenced determination; drug metabolism; chromosome 22;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200218638-A2.
XX
PD 07-MAR-2002.
XX
PF 27-AUG-2001; 2001WO-IB01544.
XX
PR 30-AUG-2000; 2000GB-0021286.
XX
PA (GEMI-) GEMINI GENOMICS PLC.
XX
PI Risinger C, Andersson MK, Lewander T, Olliasson E;
XX WPI; 2002-329785/36.
XX
PT New sequence determination oligonucleotides, useful for detecting
PT polymorphic sites in a 5' flanking region of a CYP2D6 gene, as
PT hybridization probes, as components of diagnostic assays, or in
PT ligase-based sequence determination.
XX
PS Example 3; Fig 1; 63pp; English.
XX
CC The invention relates to sequence determination oligonucleotides for
CC detecting polymorphic sites in a 5' flanking region of cytochrome P450
CC 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many
CC different xenobiotics. Human CYP2D6 gene is located on chromosome 22.
CC The oligonucleotides may be used as in situ hybridisation probes, in
CC ligase-based sequenced determination, as components of diagnostic assays,
CC as probes in sequence determination methods based on mismatches, as
CC hybridisation-based diagnostic assays, and as components of diagnostic
CC microarray. CYP2D6 is useful to predict variations in an individual's
CC ability to metabolise certain drugs. The present sequence is human
CC CYP2D6 gene.
XX
SQ Sequence 9432 BP; 1964 A; 2647 C; 2976 G; 1845 T; 0 other;

Query Match          99.0%; Score 40.6; DB 24; Length 9432;
Best Local Similarity 97.6%; Pred. No. 7.7e-07;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGAAATGTGTCYCTAAGTGTGAGTGTGAGTCT 41
   ||||||||||||||||||||||||||||||||||||||||
Db 860 GTGTGAGAGAGAAATGTGTCCTAAGTGTGAGTGTGAGTCT 900
```



DE Human nervous system related polynucleotide SEQ ID NO 13071.  
XX  
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antiscikling; antianemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX  
OS Homo sapiens.  
XX  
XX WO200159063-A2.  
XX  
XX  
XX  
XX  
XX 16-AUG-2001.  
XX  
XX  
XX 17-JAN-2001; 2001WO-US01334.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 23-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0228287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.  
XX 08-SEP-2000; 2000US-0231243.  
XX 08-SEP-2000; 2000US-0231244.  
XX 08-SEP-2000; 2000US-0231413.  
XX 08-SEP-2000; 2000US-0231414.  
XX 08-SEP-2000; 2000US-0232080.  
XX 08-SEP-2000; 2000US-0232081.  
XX 12-SEP-2000; 2000US-0231968.  
XX 14-SEP-2000; 2000US-0232397.  
XX 14-SEP-2000; 2000US-0232398.  
XX 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-02332400.  
PR 14-SEP-2000; 2000US-02332401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246533.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.



```
PN EP1033401-A2.
XX
PD
XX 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-02000610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 26285; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 248 BP; 71 A; 48 C; 95 G; 34 T; 0 other;

Query Match 56.1%; Score 23; DB 21; Length 248;
Best Local Similarity 70.7%; Pred. No. 5.7;
Matches 29; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTCGTCCTAAGTGCAGTGCAGTCT 41
Db 169 GTTTGAGAGAGAAAGTGGGCCAGGGGCCAATGCGAGTAT 209

RESULT 12
AAK85590
ID AAK85590 standard; DNA; 75384 BP.
XX
AC AAK85590;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:40402.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
```



XX AAA43426 to AAA45925 represent specifically claimed secreted expressed  
 CC sequence tags (sESTs), isolated from human, mouse, chicken and rat  
 CC tissue sources. The sESTs can have a range of activities depending on  
 CC the tissues they were isolated from. The activities include:  
 CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
 CC chemokinetic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
 CC cytostatic; antibacterial; antifungal; antiviral; antidiabetic;  
 CC antiasthmatic; vulnery; antiulcer; osteopathic; neuroprotective;  
 CC neotropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
 CC anticonvulsant; and antidepressant. The sESTs can be used for gene  
 CC therapy and in vaccines. The sESTs are useful as probes for the  
 CC identification and isolation of full-length cDNAs and genomic DNA  
 CC molecules which correspond to the sESTs. Proteins encoded by the sESTs  
 CC are useful in assays for determining biological activity and raising  
 CC antibodies. They may be useful for treatment of autoimmune disorders  
 CC (multiple sclerosis, insulin dependent diabetes), allergic conditions  
 CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
 CC osteoporosis, osteoarthritis, central nervous system disorders  
 CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
 CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
 CC disease), tumours, bacterial, fungal or viral infections, depression and  
 CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
 CC in the exemplification of the present invention.

XX SQ Sequence 468 BP; 136 A; 106 C; 91 G; 135 T; 0 other;

Query Match 53.7%; Score 22; DB 21; Length 468;  
 Best Local Similarity 70.0%; Pred. No. 17;  
 Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2 TGTGAGAGAGATGTGCYCTAAGTGTGAGTGTGAGTCT 41  
 | ||||| |||| ||| ||| ||||| |||| ||||  
 Db 335 TTGTGAGAGAGTGTGCGCTAAACACACAGTGGGAGACT 296

RESULT 14  
 AAT91307/c  
 ID AAT91307 standard; cDNA; 479 BP.  
 XX AC AAT91307;  
 XX DT 19-MAY-1998 (first entry)  
 XX DE Human H83-22 secreted protein cDNA 5' end.  
 XX KW Secreted protein; anti-inflammatory; immune stimulant; suppressant;  
 KW human; cytokine; autoimmune disease; regulator; activin; inhibin;  
 KW H83-22; ds.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FH CDS 138..479  
 FT /\*tag= a  
 FT /product= H83-22  
 FT /note= "partial secreted protein coding sequence"  
 XX PN WO9740151-A2.  
 XX PD 30-OCT-1997.  
 XX PF 14-APR-1997; 97WO-US06125.  
 XX PR 10-JAN-1997; 97US-0781226.  
 PR 19-APR-1996; 96US-0635311.  
 XX PA (GEMY ) GENETICS INST INC.  
 XX PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;  
 PI Spaulding V;  
 XX WPI; 1997-535835/49.

DR P-PSDB; AAW27295.  
 XX Nucleic acids encoding novel secreted proteins - useful e.g. as an  
 PT anti-inflammatory, immune stimulant or suppressor, etc.  
 XX Claim 23; Page 53; 81pp; English.  
 XX This cDNA sequence encodes the 5' end of a novel secreted protein,  
 CC H83-22, which is isolated from a clone, H83-22 (ATCC 98028), derived from  
 CC a human PBMC cDNA library. The nucleic acid and encoded H83-22 secreted  
 CC protein can be used for research purposes (as markers for tissues,  
 CC molecular weight markers for gels, primers, probes, etc.), for nutrition  
 CC (as C, N or carbohydrate source), as a cytokine for cell proliferation  
 CC and differentiation activity, as immune stimulants or suppressors e.g.  
 CC for viral, bacterial or fungal infections, for autoimmune diseases such  
 CC as multiple sclerosis or systemic lupus erythematosus, to regulate  
 CC haematopoiesis, for tissue growth, as an activin or inhibin, or having  
 CC chemotactic, chemokinetic, haemostatic and thrombolytic, receptor/ligand,  
 CC anti-inflammatory or tumour inhibitory activities.  
 XX SQ Sequence 479 BP; 139 A; 107 C; 92 G; 138 T; 3 other;

Query Match 53.7%; Score 22; DB 18; Length 479;  
 Best Local Similarity 70.0%; Pred. No. 17;  
 Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2 TGTGAGAGAGATGTGCYCTAAGTGTGAGTGTGAGTCT 41  
 | ||||| |||| |||| ||| ||||| |||| ||||  
 Db 352 TTGTGAGAGAGTGTGCGCTAAACACACAGTGGGAGACT 313

RESULT 15  
 AAV00431/c  
 ID AAV00431 standard; cDNA; 479 BP.  
 XX AC AAV00431;  
 XX DT 12-MAY-1998 (first entry)  
 XX DE 5' fragment of clone H83\_22.  
 XX KW Human; secreted protein; molecular weight marker; genetic fingerprinting;  
 KW antibody production; nutritional supplement; therapy; clone H83\_22;  
 KW peripheral blood mononucleocyte; PBMC; ds.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FH CDS 138..479  
 FT /\*tag= a  
 FT /transl\_except= (pos:204..206, aa:Xaa)  
 FT /transl\_except= (pos:450..452, aa:Xaa)  
 FT /note= "Xaa= unknown; no stop codon given"  
 XX PN WO9740069-A2.  
 XX PD 30-OCT-1997.  
 XX PF 14-APR-1997; 97WO-US06134.  
 XX PR 19-APR-1996; 96US-0635311.  
 XX PA (GEMY ) GENETICS INST INC.  
 XX PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;  
 PI Spaulding V;  
 XX WPI; 1997-535776/49.  
 XX P-PSDB; AAW36954.  
 XX Isolated nucleic acid clones from ATCC 98028 encode novel secreted  
 PT proteins - having many potential uses, e.g. as immunomodulators,  
 PT cell proliferation or differentiation inhibitors or haematopoiesis



PT regulators

XX  
PS Claim 44; Page 81; 114pp: English.

XX  
CC This sequence is the 5' portion of clone H83\_22, which is a  
CC polynucleotide of the invention. This sequence was isolated from a human  
CC peripheral blood mononucleocyte (PBMC) cDNA library. The polynucleotide,  
CC which encodes a secreted protein, can be used, e.g. as a tissue or  
CC molecular weight marker, in genetic fingerprinting, to raise anti-protein  
CC or anti-DNA antibodies and in interaction trap assays. The protein can be  
CC used to assay biological activity, raise antibodies for use in  
CC immunoassays, as a marker, to identify inhibitors of its interactions and  
CC as a nutritional supplement. It may also have a very wide range of  
CC therapeutic and biological activities (no examples are given to support  
CC this), e.g. cytokine or modulator of cell proliferation and  
CC differentiation, immunostimulant or immunosuppressant, haematopoiesis  
CC regulator, bone, cartilage, tendon, ligament and/or nerve tissue growth  
CC stimulator, follicle inhibitor/stimulator, chemotactic/chemokinetic,  
CC haemostatic, thrombolytic or anti-inflammatory agent, antimicrobial,  
CC biorhythm, metabolism or behaviour modifier, anti-depressant or analgesic  
CC or psoriasis treatment.

XX  
SQ Sequence 479 BP; 139 A; 107 C; 92 G; 138 T; 3 other;

Query Match 53.7%; Score 22; DB 18; Length 479;  
Best Local Similarity 70.0%; Pred. No. 17;  
Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

OY 2 TGTGAGAGAGATGTGTCYCTAAGTGTGAGTGTGAGTCT 41

DB 352 TTTGAGAGAGAGTGTGCGCCTAAACACAGTGGGAGACT 313

Search completed: February 11, 2003, 02:05:52  
Job time : 27.6596 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:51:09 ; Search time 118.283 Seconds  
(without alignments)  
5613.764 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_860\_900  
Perfect score: 41  
Sequence: 1 gctgagagagaatgtgc.....ctaagtgcagtgtgagtct 41

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	95.1	1180	14	BQ067695
2	26.8	65.4	505	9	AL723616
3	26.8	65.4	517	13	BI867734
4	26.8	65.4	595	9	AL723525
5	26.8	65.4	671	12	BF156472
6	25.2	61.5	358	17	AQ069820

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

c	7	25.2	61.5	441	10	BB690340
	8	25	61.0	1003	17	CNS028NG
	9	24.8	60.5	422	10	BB788358
c	10	24.8	60.5	823	13	BI734462
	11	24.8	60.5	1101	17	CNS00DZ4
	12	24.6	60.0	228	12	BE706013
	13	24.2	59.0	1329	12	BC289689
c	14	24.2	59.0	1667	12	BE892091
c	15	23.8	58.0	434	17	AQ817531
	16	23.8	58.0	442	10	BB728341
	17	23.6	57.6	555	9	AI942714
	18	23.6	57.6	295	9	AA302898
	19	23.6	57.6	350	17	A2841808
	20	23.6	57.6	419	13	BI881540
	21	23.6	57.6	474	10	BB694186
c	22	23.6	57.6	541	17	AQ567360
	23	23.6	57.6	544	10	BB694331
c	24	23.6	57.6	579	9	AL722438
	25	23.6	57.6	587	9	AL722350
c	26	23.6	57.6	595	17	GG0000413
	27	23.6	57.6	664	13	BI887045
c	28	23.6	57.6	701	13	BI275605
c	29	23.6	57.6	1069	17	CNS03034
	30	23.4	57.1	238	17	A2631409
c	31	23.4	57.1	361	17	A2404266
	32	23.4	57.1	852	17	CNS02J5D
	33	23.2	56.6	266	9	AI575792
c	34	23.2	56.6	274	9	AV066040
	35	23.2	56.6	404	17	A2430586
c	36	23.2	56.6	437	13	BG951973
	37	23.2	56.6	536	17	FR004226
c	38	23	56.1	230	17	A2294343
	39	23	56.1	282	17	FR0043719
	40	23	56.1	288	17	BI118260
c	41	23	56.1	294	17	A2315242
c	42	23	56.1	328	17	CNS041Z2
	43	23	56.1	361	9	AI694875
c	44	23	56.1	365	10	AW845982
	45	23	56.1	402	17	AQ762909

ALIGNMENTS

RESULT 1  
BQ067695  
LOCUS BQ067695 1180 bp mRNA linear EST 02-APR-2002  
DEFINITION AGENCOURT\_6643098 NIH\_MCC\_121 Homo sapiens CDNA clone IMAGE:5768406  
5', mRNA sequence.  
ACCESSION BQ067695  
VERSION BQ067695.1 GI:19896741  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1180)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
plate: LLAM12828 row: d column: 07  
High quality sequence start: 3  
High quality sequence stop: 536.  
Location/Qualifiers

```

source
1..1180
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5768406"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."
BASE COUNT      254 a   371 c   347 g   208 t
ORIGIN

Query Match      95.1%; Score 39; DB 14; Length 1180;
Best Local Similarity 95.1%; Pred. No. 0.00061;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCGCYCTAAGTCTCAGTGTGAGTCT 41
|||||
Db 343 GTGTGAGAGAGATGTGTCGCCCTGAGTGTGAGTGTGAGTCT 383
|||||

RESULT 2
AL723616/c
LOCUS
DEFINITION
AL723616 Danio rerio embryonic inner ear subtracted cDNA Danio
rerio cDNA clone BN0AA0672G01 5', mRNA sequence.
ACCESSION
AL723616
VERSION
AL723616.1 GI:20188220
KEYWORDS
zebrafish.
SOURCE
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 505)
AUTHORS
Coimbra,R., Weil,D., Brottier,P., Blanchard,S., Levi,M., Hardelin
,J.P., Weissenbach,J. and Petit,C.
TITLE
A subtracted cDNA library from the zebrafish (Danio rerio)
embryonic inner ear
JOURNAL
Unpublished (2002)
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..505
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="BN0AA0672G01"
/cdna_lib="Danio rerio embryonic inner ear subtracted
cDNA"
/tissue_type="inner ear"
/dev_stage="embryonic"
/note="subtracted cDNA library"
BASE COUNT      104 a   122 c   106 g   173 t
ORIGIN

Query Match      65.4%; Score 26.8; DB 9; Length 505;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 28; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAGTCTCAG 32
|||||
Db 120 GTGTGAGAGAGATGTGTCCTGAGTGTGAG 89
|||||

```

```

RESULT 3
BI867734
LOCUS
DEFINITION
BI867734 517 bp mRNA linear EST 11-OCT-2001
ft63a05.y1 Gong zebrafish ovary Danio rerio cDNA clone 5157800 5',
similar to TR:Q06054 Q06054 ZINC FINGER PROTEIN 40., mRNA
sequence.
ACCESSION
BI867734
VERSION
BI867734.1 GI:16060867
KEYWORDS
zebrafish.
SOURCE
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 517)
AUTHORS
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,K., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
TITLE
Wasnu Zebrafish EST Project 1998
JOURNAL
Unpublished (1998)
COMMENT
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
The library was constructed by Dr. Z. Gong. DNA Sequencing by:
Washington University Genome Sequencing Center St. Louis. Please
contact Zhiyuan Gong for further information on this library
(National University of Singapore, Department of Biological
Sciences, Lower Kent Ridge Road, Singapore 119260).
Seq primer: T3 ET from Amersham
High quality sequence stop: 399.
Location/Qualifiers
1..517
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5157800"
/cdna_lib="Gong zebrafish ovary"
/sex="female"
/dev_stage="4-5 month"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary (pooled); Vector: pBluescript SK-;
Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolated from
the ovaries of 2 female adult zebrafish (4-5 month old).
cDNAs were made using oligo-dT primers and inserted into
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
mass-excised to pBluescript SK- following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Zhiyuan Gong for further information on
this library (National University of Singapore,
Department of Biological Sciences, Lower Kent Ridge Road,
Singapore 119260)."
BASE COUNT      181 a   100 c   118 g   118 t
ORIGIN

Query Match      65.4%; Score 26.8; DB 13; Length 517;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 28; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAGTCTCAG 32
|||||
Db 415 GTGTGAGAGAGATGTGTCCTGAGTGTGAG 446
|||||

RESULT 4
AL723525
LOCUS
DEFINITION
AL723525 595 bp mRNA linear EST 18-APR-2002
Danio rerio embryonic inner ear subtracted cDNA Danio

```

Accession AL723525 Version AL723525.1 Keywords EST. Source zebrafish. Organism Danio rerio Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio. Reference 1 (bases 1 to 595) Authors Coimbra, R., Weil, D., Brottier, P., Blanchard, S., Levi, M., Hardelin, J. P., Weissenbach, J. and Petit, C. Title A subtracted cDNA library from the zebrafish (Danio rerio) embryonic inner ear Journal Unpublished (2002) Comment Contact: Genoscope Genoscope - Centre National de Sequencage Bp 191 91006 Evry cedex - France Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr.	Consortium/LLNL, send email to: info@image.llnl.gov Seq primer: T3 ET from Amersham High quality sequence stop: 505. Features Location/Qualifiers 1..671 /organism="Danio rerio" /strain="AB" /db_xref="taxon:7955" /clone_lib="3817540" /clone_lib="Sugano Kawakami zebrafish DRA" /sex="mixed (one male and one female, including unfertilized eggs)" /dev_stage="adult" /lab_host="DH10B (phage resistant)" /note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCTTTTCTTTTCTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [GTGTGGCCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science) and kindly donated by Dr. Koichi Kawakami. Custom primers for sequencing: 5' end primer CTTCGTCTCTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACA."
Base Count Origin Query Match Best Local Similarity Matches 28; Conservative 1; Mismatches 3; Indels 0; Gaps 0;	Base Count Origin Query Match Best Local Similarity Matches 28; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GTGTGAGAGAGATGTGTCYCTAAGTGTGAC 32       Db 477 GTGTGAGAGAGATGTGTCGTCTGAGTGTGAG 508 	Qy 1 GTGTGAGAGAGATGTGTCYCTAAGTGTGAC 32       Db 496 GTGTGAGAGAGATGTGTCGTCTGAGTGTGAG 527 
Result 5 BF156472 Locus Definition f152g03.y1 Sugano Kawakami zebrafish DRA Danio rerio cDNA clone 3817540 5', similar to TR:Q9UG02 Q9UG02 HYPOTHETICAL 80.6 KD PROTEIN ; mRNA sequence. Accession BF156472 Version BF156472.1 Keywords EST. Source zebrafish. Organism Danio rerio Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio. Reference 1 (bases 1 to 671) Authors Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R. Title WashU Zebrafish EST Project 1998 Journal Unpublished (1998) Comment Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: zbrafish@watson.wustl.edu Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution information can be found through the I.M.A.G.E.	Result 6 AQ069820/c Locus Definition HS_2253_A2_B02_MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2253 Col=4 Row=C, DNA sequence. Accession AQ069820 Version AQ069820.1 Keywords GSS. Source human. Organism Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo. Reference 1 (bases 1 to 358) Authors Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L. Title Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome Journal Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999) Medline 99380589 Comment Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Sequence Tagged Connector Plate: 2253 row: C column: 4 Class: BAC ends High quality sequence stop: 358. Features Location/Qualifiers 1..358 /organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone_plate="2253 Col-4 Row=C"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/notes="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"
BASE COUNT      134 a   62 c   42 g   117 t   3 others
ORIGIN

Query Match      61.5%; Score 25.2; DB 17; Length 358;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 30; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy  2  TGTGAGAGAAATGTGTCYCTAAGTGTGTCAGTGTGAGTCT 41
      ||| ||||| ||||| ||| ||| ||| |||||
Db  131 TGTGAGAGAAATGTGCTAAAGAGTCTGTGATGCT 92

RESULT 7
BB690340      441 bp      mRNA      linear      EST 10-OCT-2001
LOCUS      BB690340 RIKEN full-length enriched, 12 days embryo female
DEFINITION      mullerian duct Mus musculus cDNA clone 6820447I12 3', mRNA
sequence.
ACCESSION      BB690340
VERSION      BB690340.1 GI:16017073
KEYWORDS      EST.
SOURCE      house mouse
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 441)
Hayatsu.N., Hiramoto.K., Hiraoka.T., Hirozane.T., Imotani.K., Ishii
.Y., Ito.M., Kawai.J., Kojima.Y., Konno.H., Kouda.M., Matsuyama.T.,
Nakamura.M., Nishi.K., Nomura.K., Numasaki.R., Okazaki.Y., Okido.T.,
Saito.R., Sakai.C., Sakai.K., Sakazume.N., Sasaki.D., Sato.K.,
Shibata.K., Shingawa.A., Shiraki.T., Sogabe.Y., Suzuki.H., Tagawa
.A., Takahashi.F., Takaku-Akashira.S., Tanaka.T., Tomaru.A., Toya.T.,
Watahiki.A., Yasunishi.A., Muramatsu.M. and Hayashizaki.Y.
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
.M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki.Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. . 10 (10), 1617-1630 (2000)
waqi,K., Fujiwaka,S., Inoue,K., Togawa.Y., Izawa.M., Ohara.E.,
Watahiki,M., Yoneda.Y., Ishikawa.T., Ozawa.K., Tanaka,T., Matsuura
.S., Kawai,J., Okazaki.Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
10 (11), 1757-1771 (2000)
Konno.H., Fukunishi,Y., Shibata.K., Itoh.M., Carninci,P., Sugahara
.Y. and Hayashizaki.Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers

```

```

1. 441
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="6820447I12"
/clone_lib="RIKEN full-length enriched, 12 days embryo
female mullerian duct"
/sex="female"
/tissue_type="mullerian duct includes surrounding region"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGATCCAGAGCTCTTTTITTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATTCGAGTCTAATTAATTAATCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC 1: Cloning sites, 5' end: SalI; 3' end: BamHI"
BASE COUNT      103 a   112 c   89 g   137 t
ORIGIN

Query Match      61.5%; Score 25.2; DB 10; Length 441;
Best Local Similarity 78.9%; Pred. No. 43;
Matches 30; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy  1  GTGTGAGAGAAATGTGTCYCTAAGTGTGTCAGTGTGAG 38
      ||| ||||| ||||| ||| ||| ||| |||||
Db  229 GTGAGAGAGATGTGTGCGGTGTGTGTGTGTGTGTGAG 266

RESULT 8
CNS028NG/c
LOCUS      CNS028NG 1003 bp      DNA      linear      GSS 12-MAY-2000
DEFINITION      Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
245K23 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION      AL186181
VERSION      AL186181.1 GI:7824285
KEYWORDS      GSS: genome survey sequence.
SOURCE      Tetraodon nigroviridis.
ORGANISM      Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 1003)
Roest-Crollius,H., Jaillon,O., Basilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 1003)
Roest-Crollius,H., Jaillon,O., Basilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
3 (bases 1 to 1003)
Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (12-APR-2000)
REFERENCE      This sequence is a single read and was generated as part of a large
COMMENT

```



```
/lab_host="DH10B (phage-resistant)"
/notes="organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
BASE COUNT      178 a   183 c   252 g   210 t
ORIGIN

Query Match      60.5%; Score 24.8; DB 13; Length 823;
Best Local Similarity 76.3%; Pred. No. 85;
Matches 29; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGAAATGTGTCGYCTAAGTCAGTGTGAG 38
||||| ||||| || ||||| : ||||| ||||| |||||
Db 459 GTGTGAGGAGAGAGCGCTGCTGTGTGTGTGTGTGAG 496

RESULT 11
CNS00D24/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR28C03 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION      AL077425
VERSION        AL077425.1 GI:4948682
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org the BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
Source
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR28C03"
/clone_lib="RPCI-98"
/notes="end : TET3"
BASE COUNT      338 a   404 c   85 g   142 t   132 others
ORIGIN

Query Match      60.5%; Score 24.8; DB 17; Length 1101;
Best Local Similarity 74.4%; Pred. No. 99;
Matches 29; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGAAATGTGTCGYCTAAGTCAGTGTGAGT 39
||||| ||||| || ||||| : ||||| ||||| |||||
Db 701 GTGTGAGAGTGAATGTGTGTCNGGGTGTGWTGTGTGT 663

RESULT 12
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 228)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
M.J., Soares,F., Brentani,R.K., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=4t2-MR0-HT0158-270
600-008-d04&t3=2000-06-27&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 20
High quality sequence stop: 167.
Location/Qualifiers
1..228
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0158"
/dev_stage="Adult"
/notes="organ: head_neck; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      32 a   24 c   81 g   91 t
ORIGIN

Query Match      60.0%; Score 24.6; DB 12; Length 228;
Best Local Similarity 73.2%; Pred. No. 49;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGAAATGTGTCGYCTAAGTCAGTGTGAGTCT 41
||||| ||||| || ||||| : ||||| ||||| |||||
Db 170 GAGTCTGAGTGAATGTGTGGTCTCTGTGTGTGTGAGTCT 210

RESULT 13
BG289689
LOCUS
DEFINITION
602384668P1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4513601 5',
mRNA sequence.
ACCESSION      BG289689
VERSION        BG289689.1 GI:13045729
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```





Qy 1 GTGTGAGAGAGAACTGTGTGTCYCTAAGTCTCAGTGTGAG 38  
||||| :| |||||  
Db 383 GTGTGNCATGAATCTGTGGTGTGGGTCTCAGTGTGTG 346

Search completed: February 11, 2003, 03:31:19  
Job time : 123.283 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 21:12:29 ; Search time 3.11557 Seconds  
(without alignments)  
4035.781 Million cell updates/sec

Title: US-09-942-310-2\_Copy\_860\_900

Perfect score: 41

Sequence: 1 ggtgagagagaagtgtgc.....ctaagtgtcagtgtagtct 41

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5b\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	53.7	8310	3	US-08-870-126-11 Sequence 11, Appl
2	22	53.7	8310	4	US-09-445-247-11 Sequence 11, Appl
3	22	53.7	14985	1	US-08-652-972A-6 Sequence 6, Appl
4	22	53.7	14985	5	PCT-US96-06231A-6 Sequence 6, Appl
5	21	51.2	3958	1	US-08-435-933-5 Sequence 5, Appl
6	21	51.2	3958	5	PCT-US96-06035-5 Sequence 5, Appl
c 7	20.4	49.8	1194	4	US-09-134-001C-2256 Sequence 2256, Ap
8	20.4	49.8	16442	3	US-08-781-891-208 Sequence 208, App
9	20.4	49.8	111282	4	US-09-754-250-3 Sequence 3, Appl
10	20.2	49.3	350	1	US-08-472-217-4 Sequence 4, Appl
11	20.2	49.3	350	3	US-08-760-534A-4 Sequence 4, Appl
12	20.2	49.3	2196	1	US-08-472-217-3 Sequence 3, Appl
13	20.2	49.3	2196	3	US-08-760-534A-3 Sequence 3, Appl
c 14	19.8	48.3	62	1	US-08-222-177A-56 Sequence 56, Appl
15	19.8	48.3	1493	1	US-08-340-820-24 Sequence 24, Appl
16	19.8	48.3	1493	1	US-08-593-535-24 Sequence 24, Appl
c 17	19.8	48.3	2759	4	US-09-144-367-1 Sequence 1, Appl
18	19.8	48.3	4185	3	US-08-964-268-1 Sequence 1, Appl
c 19	19.8	48.3	6953	1	US-07-805-123C-2 Sequence 2, Appl
c 20	19.8	48.3	6953	1	US-08-033-081B-2 Sequence 3, Appl
c 21	19.8	48.3	50000	4	US-09-146-053-3 Sequence 3, Appl
22	19.6	47.8	2065	1	US-08-261-822A-11 Sequence 11, Appl
23	19.6	47.8	2065	5	PCT-US95-07744A-11 Sequence 11, Appl
24	19.4	47.3	11613	1	US-08-484-044-10 Sequence 10, Appl
25	19.2	46.8	1977	4	US-09-595-424-5 Sequence 5, Appl
26	19.2	46.8	2294	4	US-09-643-597-123 Sequence 123, App
27	19.2	46.8	3258	4	US-09-595-424-7 Sequence 7, Appl

28	18.8	45.9	270	1	US-08-222-177A-51 Sequence 51, Appl
29	18.8	45.9	343	1	US-08-320-559-36 Sequence 36, Appl
30	18.8	45.9	343	3	US-08-545-860D-36 Sequence 36, Appl
31	18.8	45.9	343	5	PCT-US94-04496-36 Sequence 36, Appl
c 32	18.6	45.4	2997	3	US-08-486-343A-3 Sequence 3, Appl
c 33	18.6	45.4	2997	5	PCT-US95-07349-3 Sequence 3, Appl
34	18.6	45.4	11873	2	US-08-970-269A-32 Sequence 32, Appl
35	18.6	45.4	11873	4	US-09-407-562-32 Sequence 32, Appl
36	18.6	45.4	11878	2	US-08-970-269A-31 Sequence 31, Appl
37	18.6	45.4	11878	4	US-09-407-562-31 Sequence 31, Appl
38	18.6	45.4	11883	2	US-08-970-269A-28 Sequence 28, Appl
39	18.6	45.4	11883	4	US-09-407-562-28 Sequence 28, Appl
c 40	18.4	44.9	3172	1	US-07-741-940-3 Sequence 3, Appl
c 41	18.4	44.9	3172	1	US-08-289-548A-3 Sequence 3, Appl
c 42	18.4	44.9	3172	1	US-08-452-655B-3 Sequence 3, Appl
c 43	18.4	44.9	3172	3	US-08-450-582-3 Sequence 3, Appl
c 44	18.4	44.9	3172	3	US-08-450-582-3 Sequence 3, Appl
c 45	18.4	44.9	3172	4	US-08-449-731-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1  
US-08-870-126-11  
; Sequence 11, Application US/08870126  
; Patent No. 6048702  
; GENERAL INFORMATION:  
; APPLICANT: Prendergast, George C.  
; APPLICANT: Sakamuro, Daitoku  
; TITLE OF INVENTION: Murine and Human Box-Dependent  
; TITLE OF INVENTION: NYC-Interacting Protein (Binl) and Uses Therefor  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P O Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/870,126  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/435,454  
; FILING DATE: 05-MAY-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/652,972  
; FILING DATE: 24-MAY-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kodroff, Cathy A.  
; REGISTRATION NUMBER: 33,980  
; REFERENCE/DOCKET NUMBER: WST60CUSA  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-540-9200  
; TELEFAX: 215-540-5818  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 8310 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 680..765 /note= "exon 7"  
; OTHER INFORMATION:

```
;
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1052..1127
; OTHER INFORMATION: /note= "exon 8"
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2503..2585
; OTHER INFORMATION: /note= "exon 9"
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4059..4103
; OTHER INFORMATION: /note= "exon 10"
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5543..5687
; OTHER INFORMATION: /note= "exon 11"
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5688..5929
; OTHER INFORMATION: /note= "putative.alt.exon"
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 7094..7221
; OTHER INFORMATION: /note= "exon 12"
;
; US-08-870-126-11
;
; Query Match 53.7%; Score 22; DB 3; Length 8310;
; Best Local Similarity 73.7%; Pred. No. 4.2;
; Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
;
; QY 1 GTGTGAGAGAGAAATGTGTGTCYCTAAGTGTGTCAGTGTGAG 38
; ||||| ||||| ||||| ||||| ||||| ||||| |||||
; Db 3220 GTGTGAGGGGACTGTGTGTGACAGGTGTAAGTGTGTG 3257
```

```
RESULT 2
US-09-445-247-11
; Sequence 11, Application US/09445247
; Patent No. 6410238
; GENERAL INFORMATION:
; APPLICANT: Wistar Institute of Anatomy & Biology
; Prendergast, George C.
; Sakamuro, Daitoku
; TITLE OF INVENTION: Box-Dependent MYC-Interacting Protein
; (Bin1) Compositions and Uses Therefor
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/445,247
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/870,126
; FILING DATE: 06-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST60DPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 11:
```

```
;
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: exon
; LOCATION: 680..765
; OTHER INFORMATION: /note= "exon 7"
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1052..1127
; OTHER INFORMATION: /note= "exon 8"
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2503..2585
; OTHER INFORMATION: /note= "exon 9"
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4059..4103
; OTHER INFORMATION: /note= "exon 10"
;
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5543..5687
; OTHER INFORMATION: /note= "exon 11"
;
; NAME/KEY: exon
; LOCATION: 7093..7221
; OTHER INFORMATION: /note= "exon 12A"
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-445-247-11
;
; Query Match 53.7%; Score 22; DB 4; Length 8310;
; Best Local Similarity 73.7%; Pred. No. 4.2;
; Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
;
; QY 1 GTGTGAGAGAGAAATGTGTGTCYCTAAGTGTGTCAGTGTGAG 38
; ||||| ||||| ||||| ||||| ||||| ||||| |||||
; Db 3220 GTGTGAGGGGACTGTGTGTGACAGGTGTAAGTGTGTG 3257
;
; RESULT 3
US-08-652-972A-6
; Sequence 6, Application US/08652972A
; Patent No. 5723581
; GENERAL INFORMATION:
; APPLICANT: Prendergast, George C.
; APPLICANT: Sakamuro, Daitoku
; TITLE OF INVENTION: Murine and Human Box-Dependent
; TITLE OF INVENTION: MYC-Interacting Protein (Bin1) and Uses Therefor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, P O Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/652,972A
; FILING DATE: 24-MAY-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,454
; FILING DATE: 05-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
```

REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60BUSA  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14985 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 1332  
OTHER INFORMATION: /note= "unsequenced segment"  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 3225  
OTHER INFORMATION: /note= "unsequenced segment"  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 7209  
OTHER INFORMATION: /note= "unsequenced segment"  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 11097  
OTHER INFORMATION: /note= "unsequenced segment"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1..324  
OTHER INFORMATION: /note= "Exon 1"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 325..1618  
OTHER INFORMATION: /note= "Exon 2"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1619..3174  
OTHER INFORMATION: /note= "Exon 3"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 3175..4365  
OTHER INFORMATION: /note= "Exon 4"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4441..11518  
OTHER INFORMATION: /note= "Exon 5"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 11519..11850  
OTHER INFORMATION: /note= "Exon 6"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 11851..12240  
OTHER INFORMATION: /note= "Exon 7"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 12241..14129  
OTHER INFORMATION: /note= "Exon 8"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 14130..14985  
OTHER INFORMATION: /note= "Exon 9"

US-08-652-972A-6

Query Match 53.7%; Score 22; DB 1; Length 14985;  
Best Local Similarity 73.7%; Pred. No. 4.8;  
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAGTGTCAAGTGTGAG 38

||||||| | | | | | | | | | | | | | | | |

DB 2336 GTGTGAGGGGGACTGTGTGTGACAGGTGTAAGTGTGTG 2373

RESULT 4  
PCT-US96-06231A-6  
Sequence 6, Application PC/TUS9606231A  
GENERAL INFORMATION:  
APPLICANT: Wistar Institute of Anatomy & Biology  
TITLE OF INVENTION: Murine and Human Box-Dependent  
TITLE OF INVENTION: Myc-Interacting Protein (BIN1) and Uses Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P O Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/06231A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,454  
FILING DATE: 05-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST60APCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14985 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 1332  
OTHER INFORMATION: /note= "unsequenced segment"  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 3225  
OTHER INFORMATION: /note= "unsequenced segment"  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 7209  
OTHER INFORMATION: /note= "unsequenced segment"  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 11097  
OTHER INFORMATION: /note= "unsequenced segment"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1..324  
OTHER INFORMATION: /note= "Exon 1"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 325..1618  
OTHER INFORMATION: /note= "Exon 2"  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1619..3174  
OTHER INFORMATION: /note= "Exon 3"  
FEATURE:  
NAME/KEY: exon

```
; LOCATION: 3175..4365
; OTHER INFORMATION: /note= "Exon 4"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4441..11518
; OTHER INFORMATION: /note= "Exon 5"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 11519..11850
; OTHER INFORMATION: /note= "Exon 6"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 11851..12240
; OTHER INFORMATION: /note= "Exon 7"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12241..14129
; OTHER INFORMATION: /note= "Exon 8"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14130..14985
; OTHER INFORMATION: /note= "Exon 9"
; PCT-US96-06231A-6
;
Query Match 53.7%; Score 22; DB 5; Length 14985;
Best Local Similarity 73.7%; Pred. No. 4.8;
Matches 28; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAAGTGCAGTGTGAG 38
||||| ||| ||||| | |||| ||||| |
Db 2336 GTGTGAGGGGAGTGTGTGACAGAGTGAAGTGTGTG 2373

RESULT 5
US-08-435-933-5
; Sequence 5, Application US/08435933
; Patent No. 5693492
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; APPLICANT: Arena, Joseph P.
; APPLICANT: Pares, Philip S.
; APPLICANT: Liu, Ken K.
; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
; TITLE OF INVENTION: CHANNELS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: 126 East Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/435,933
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen, III John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19264
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; LOCATION: 3175..4365
; OTHER INFORMATION: /note= "Exon 4"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4441..11518
; OTHER INFORMATION: /note= "Exon 5"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 11519..11850
; OTHER INFORMATION: /note= "Exon 6"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 11851..12240
; OTHER INFORMATION: /note= "Exon 7"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 12241..14129
; OTHER INFORMATION: /note= "Exon 8"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 14130..14985
; OTHER INFORMATION: /note= "Exon 9"
; PCT-US96-06231A-6
;
Query Match 51.2%; Score 21; DB 5; Length 3958;
Best Local Similarity 69.2%; Pred. No. 9.1;
Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAAGTGCAGTGTGAGT 39
||||| ||| ||||| | |||| ||||| |||
Db 114 GTGTGTGTGAGTGTTTTGTACATGTGCCAGTGTGAGT 152

RESULT 6
PCT-US96-06035-5
; Sequence 5, Application PC/TUS9606035
; GENERAL INFORMATION:
; APPLICANT: Cully, Doris F.
; APPLICANT: Arena, Joseph P.
; APPLICANT: Pares, Philip S.
; APPLICANT: Liu, Ken K.
; TITLE OF INVENTION: DNA ENCODING GLUTAMATE GATED CHLORIDE
; TITLE OF INVENTION: CHANNELS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jody M. Giesser
; STREET: 126 East Lincoln Avenue - P.O. BOX 2000-0907
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/06035
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Giesser, Jody M.
; REGISTRATION NUMBER: 32,838
; REFERENCE/DOCKET NUMBER: 19264 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3046
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; PCT-US96-06035-5
;
Query Match 51.2%; Score 21; DB 5; Length 3958;
Best Local Similarity 69.2%; Pred. No. 9.1;
Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAAGTGCAGTGTGAGT 39
||||| ||| ||||| | |||| ||||| |||
Db 114 GTGTGTGTGAGTGTTTTGTACATGTGCCAGTGTGAGT 152

RESULT 7
US-09-134-001C-2256/c
; Sequence 2256, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
```

FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2256  
LENGTH: 1194  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2256

Query Match 49.8%; Score 20.4; DB 4; Length 1194;  
Best Local Similarity 75.0%; Pred. No. 12;  
Matches 24; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 10 AGAATGTGGCYCTAAGTCAGTCAGTGACTCT 41  
| | | | | | | | | | | | | | | | | | | | | |  
Db 267 AAAATGTGCCATAAATAATCATGATTATTTT 236

RESULT 8  
US-08-781-891-208  
Sequence 208, Application US/08781891  
Patent No. 6090620  
GENERAL INFORMATION:  
APPLICANT: Fu, Ying-Hui  
APPLICANT: Yu, Chang-En  
APPLICANT: Oshima, Junko  
APPLICANT: Mulligan, John T.  
APPLICANT: Schellenberg, Gerald D.  
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO  
TITLE OF INVENTION: WERNER'S SYNDROME  
NUMBER OF SEQUENCES: 209  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/781,891  
FILING DATE: 27-DEC-1996  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 6090620tenburg Ph.D., Carol  
REGISTRATION NUMBER: 39,317  
REFERENCE/DOCKET NUMBER: 240052.419  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 208:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16442 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-781-891-208

Query Match 49.8%; Score 20.4; DB 3; Length 16442;  
Best Local Similarity 80.0%; Pred. No. 23;  
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAGTGTC 30  
| | | | | | | | | | | | | | | | | | | | | |

Db 5570 GTCTGAGATAAAATGTAAGCACTAAGTGTC 5599

RESULT 9  
US-09-754-250-3  
Sequence 3, Application US/09754250  
Patent No. 6376225  
GENERAL INFORMATION:  
APPLICANT: Wei, Ming-Hui et al  
TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE  
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF  
FILE REFERENCE: CL001063  
CURRENT APPLICATION NUMBER: US/09/754,250  
CURRENT FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 111282  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(111282)  
OTHER INFORMATION: n = A,T,C or G  
US-09-754-250-3

Query Match 49.8%; Score 20.4; DB 4; Length 111282;  
Best Local Similarity 71.1%; Pred. No. 36;  
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAGTCAGTGTCGAG 38  
| | | | | | | | | | | | | | | | | | | | | |  
Db 81928 GTCTCAGTCAGAGAGGTGGGCTCAGTGAGAGGGTGAG 81965

RESULT 10  
US-08-472-217-4  
Sequence 4, Application US/08472217  
Patent No. 5726058  
GENERAL INFORMATION:  
APPLICANT: Alanen-Kurki, Leena  
APPLICANT: Auvinen, Petri  
APPLICANT: Jaakkola, Panu  
APPLICANT: Jalkanen, Markku  
APPLICANT: Lepp, Sirpa  
APPLICANT: Mali, Markku  
APPLICANT: Vihinen, Tapani  
APPLICANT: W rrl, Anni  
TITLE OF INVENTION: Syndecan Stimulation of Cellular  
TITLE OF INVENTION: Differentiation  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1100 New York Avenue, Suite 600  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,217  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/206,186  
FILING DATE: 07-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/988,427

```
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1102.0050003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-472-217-4
;
Query Match 49.3%; Score 20.2; DB 1; Length 350;
Best Local Similarity 71.4%; Pred. No. 11;
Matches 25; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 AGAGAGAATGTGTCYCTAAGTGTGAGTGTGAGTC 40
   | | | | | | | | | | | | | | | | | | | |
Db 146 AAAGAGACTGCGTGCCTTCACTGCCTGTGTGAATC 180

RESULT 11
US-08-760-534A-4
; Sequence 4, Application US/08760534A
; Patent No. 6017727
; GENERAL INFORMATION:
; APPLICANT: JALKANEN, MARKKU
; APPLICANT: JAAKKOLA, PANU
; APPLICANT: VIHINEN, TAPANI
; TITLE OF INVENTION: SYNDECAN ENHANCER ELEMENT AND SYNDECAN
; TITLE OF INVENTION: STIMULATION OF CELLULAR DIFFERENTIATION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,534A
; FILING DATE: 02-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/206,186
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FI93/00514
; FILING DATE: 01-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: CIMBALA, MICHELE A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1708.0050004/MAC
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-760-534A-4
```

```
Query Match 49.3%; Score 20.2; DB 3; Length 350;
Best Local Similarity 71.4%; Pred. No. 11;
Matches 25; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 AGAGAGAATGTGTCYCTAAGTGTGAGTGTGAGTC 40
   | | | | | | | | | | | | | | | | | | | |
Db 146 AAAGAGACTGCGTGCCTTCACTGCCTGTGTGAATC 180

RESULT 12
US-08-472-217-3
; Sequence 3, Application US/08472217
; Patent No. 5726058
; GENERAL INFORMATION:
; APPLICANT: Alanen-Kurki, Leena
; APPLICANT: Auvinen, Petri
; APPLICANT: Jaakkola, Panu
; APPLICANT: Jalkanen, Markku
; APPLICANT: Lepp, Sirpa
; APPLICANT: Mali, Markku
; APPLICANT: Vihinen, Tapani
; APPLICANT: Wrrl, Anni
; TITLE OF INVENTION: Syndecan Stimulation of Cellular
; TITLE OF INVENTION: Differentiation
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,217
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/206,186
; FILING DATE: 07-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,427
; FILING DATE: 01-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Cimbala, Michele A.
; REGISTRATION NUMBER: 33,851
; REFERENCE/DOCKET NUMBER: 1102.0050003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2196 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
US-08-472-217-3
;
Query Match 49.3%; Score 20.2; DB 1; Length 2196;
Best Local Similarity 71.4%; Pred. No. 17;
Matches 25; Conservative 1; Mismatches 9; Indels 0; Gaps 0;
```

```
QY 6 AGAGAGAATGTGTCYCTAAGTGTGAGTGTGAGTC 40
   | | | | | | | | | | | | | | | | | | | |
Db 1295 AAAGAGACTGCGTGCCTTCACTGCCTGTGTGAATC 1329
```



```

STATE: Wisconsin
COUNTRY: USA
ZIP: 53717-1914
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222.177A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/341,562
FILING DATE: 21-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Sara, Charles S.
REGISTRATION NUMBER: 30,492
REFERENCE/DOCKET NUMBER: 09865.601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 831-2100
TELEFAX: (608) 831-2106
TELEX:
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 62 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
CLONE:
IMMEDIATE SOURCE:
US-08-222-177A-56

Query Match 48.3%; Score 19.8; DB 1; Length 62;
Best Local Similarity 65.9%; Pred. No. 10;
Matches 27; Conservative 1; Mismatches 13; Indels

QY 1 GTGTGAGAGAAATGTGTGCYCTAAGTGTCACTGTGAGTCT 41
Db 47 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7

RESULT 15
US-08-340-820-24
Sequence 24, Application US/08340820
Patent No. 5512460
GENERAL INFORMATION:
APPLICANT: NARUO, Ken-ichi
APPLICANT: SEKO, Chisako
APPLICANT: KUROKAWA, Tsutomu
APPLICANT: KONDO, Tatsuya
TITLE OF INVENTION: GLIA ACTIVATING FACTOR AND ITS
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,820
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

•

Result No.	Score	Query		Length	DB	ID	Description
		Match					
c 1	22	53.7	1040	10	US-09-962-436-276	Sequence 276, Appl	
	22	53.7	1078	9	US-10-114-093-25	Sequence 25, Appl	
	3	21.8	53.2	485	10	US-09-864-761-2538	Sequence 2538, Ap
	4	21.4	52.2	659158	9	US-09-771-208-20	Sequence 20, Appl
c 5	21	51.2	339	10	US-09-867-701-8075	Sequence 8075, Ap	
	6	51.2	1540	10	US-09-393-634-8	Sequence 8, Appl	
	7	51.2	3449	9	US-09-510-332-84	Sequence 84, Appl	
	8	20.8	50.7	10351	9	US-09-874-470-5	Sequence 5, Appli
c 10	20.6	50.2	303	9	US-10-040-739-1149	Sequence 1149, Ap	
	20.6	50.2	3980	9	US-09-949-842-6	Sequence 6, Appli	
	11	20.4	49.8	2331	10	US-09-764-871-3691	Sequence 3691, Ap
	12	20.4	49.8	10828	10	US-09-942-325A-2	Sequence 2, Appli
c 13	20.4	49.8	13329	10	US-09-942-325A-1	Sequence 1, Appli	
	14	20.4	49.8	111282	12	US-10-094-989-3	Sequence 3, Appli
	15	20.2	49.3	513	10	US-09-864-761-7320	Sequence 7320, Ap
	16	20.2	49.3	1103	10	US-09-778-844-52	Sequence 52, Appl
c 17	20.2	49.3	7221	10	US-09-764-855-234	Sequence 234, App	
	18	20.2	49.3	26668	10	US-09-962-832-222	Sequence 222, App
	19	20.2	49.3	155074	9	US-10-036-188-6	Sequence 6, Appli
	20	20.2	49.3	155074	9	US-10-036-188-6	Sequence 6, Appli

```
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKeough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 1078
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-114-893-25

Query Match      53.7%; Score 22; DB 9; Length 1078;
Best Local Similarity 70.0%; Pred. No. 7.6;
Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY  2 TGTGAGAGAGATGTGTCYCTAGTGTGAGTCT 41
    | ||||| |||| ||: || ||||| ||||
Db  321 TTTGAGAGAGATGTGCGCTAAACACACAGTGGAGACT 282

RESULT 3
US-09-864-761-2538/c
; Sequence 2538, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2538
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL035665.22
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 15
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
US-09-864-761-2538

Query Match      53.2%; Score 21.8; DB 10; Length 485;
Best Local Similarity 70.7%; Pred. No. 7.5;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY  1 GTGTGAGAGAGATGTGTCYCTAAGTGCAGTGTGAGTCT 41
    | ||| | ||||| || ||||| ||||| |||||
Db  231 GCGTGTGTGTCATGTCTACTGTGTGTCAGTGTGTGTGT 191

RESULT 4
US-09-771-208-20
; Sequence 20, Application US/09771208
; Patent No. US20020155564A1
; GENERAL INFORMATION:
; APPLICANT: MEDRANO, JUAN
; APPLICANT: BRADFORD, ERIC
; APPLICANT: HORVAT, SIMON
; TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
; FILE REFERENCE: 407T-923710US
; CURRENT APPLICATION NUMBER: US/09/771,208
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 08/999,477
; PRIOR FILING DATE: 1997-12-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 659158
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (123459)..(123478)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (602466)..(602485)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (546998)..(547017)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (494715)..(494814)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (390986)..(391005)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
```

LOCATION: (346860)..(346823)  
OTHER INFORMATION: n is unidentified a, c, g, or t  
NAME/KEY: misc\_feature  
LOCATION: (317174)..(317193)  
OTHER INFORMATION: n is unidentified a, c, g, or t  
NAME/KEY: misc\_feature  
LOCATION: (280353)..(280373)  
OTHER INFORMATION: n is unidentified a, c, g, or t  
NAME/KEY: misc\_feature  
LOCATION: (271829)..(271848)  
OTHER INFORMATION: n is unidentified a, c, g, or t  
NAME/KEY: misc\_feature  
LOCATION: (183872)..(183891)  
OTHER INFORMATION: n is unidentified a, c, g, or t  
NAME/KEY: misc\_feature  
LOCATION: (170625)..(170645)  
OTHER INFORMATION: n is unidentified a, c, g, or t  
NAME/KEY: misc\_feature  
LOCATION: (132680)..(132700)  
OTHER INFORMATION: n is unidentified a, c, g, or t  
NAME/KEY: misc\_feature  
OTHER INFORMATION: n is a, c, g, or t  
US-09-771-208-20

Query Match 52.2%; Score 21.4; DB 9; Length 659158;  
Best Local Similarity 68.3%; Pred. No. 61;  
Matches 28; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAGTGTGAGTGTGAGTCT 41  
||||| | | | | | | | | | | | | | | | | | | | | |  
Db 105340 GTGTGTGTGTGTGTGTGTGTATATGTGTGAGTGTGTAT 105380

RESULT 5  
US-09-867-701-8075/C  
Sequence 8075, Application US/09867701  
Patent No. US20020132237A1  
GENERAL INFORMATION:  
APPLICANT: Aglate, Paul A.  
APPLICANT: Jones, Robert  
APPLICANT: Harlocker, Susan L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.497  
CURRENT APPLICATION NUMBER: US/09/867,701  
CURRENT FILING DATE: 2001-05-29  
NUMBER OF SEQ ID NOS: 10912  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8075  
LENGTH: 339  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-867-701-8075

Query Match 51.2%; Score 21; DB 10; Length 339;  
Best Local Similarity 69.2%; Pred. No. 14;  
Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 3 GTGAGAGAGATGTGTCYCTAAGTGTGAGTGTGAGTCT 41  
| | | | | | | | | | | | | | | | | | | | | |  
Db 57 GGGAGGAGAGAAATATGCTGTAATGTCAGGGAAGTCT 19

RESULT 6  
US-09-393-634-8  
Sequence 8, Application US/09393634  
Patent No. US20020051997A1  
GENERAL INFORMATION:  
APPLICANT: Zuker, Charles S.  
APPLICANT: Adler, Jon Elliot  
APPLICANT: Ryba, Nick  
APPLICANT: Mueller, Ken  
APPLICANT: Hoon, Mark

APPLICANT: The Regents of the University of California  
APPLICANT: The Government of the United States of America  
APPLICANT: as represented by the Secretary of the  
APPLICANT: Department of Health and Human Services  
TITLE OF INVENTION: SF, a No. US20020051997A1el Family of Taste Receptors  
FILE REFERENCE: 02307E-0980000US  
CURRENT APPLICATION NUMBER: US/09/393,634  
CURRENT FILING DATE: 1999-09-10  
NUMBER OF SEQ ID NOS: 92  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 1540  
TYPE: DNA  
ORGANISM: Rattus sp.  
FEATURE:  
OTHER INFORMATION: rat GR04 sequence approximately 1100 bp 5' to SEQ  
OTHER INFORMATION: ID NO:92  
US-09-393-634-8

Query Match 51.2%; Score 21; DB 10; Length 1540;  
Best Local Similarity 69.2%; Pred. No. 21;  
Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAGTGTGAGTGTGAGT 39  
||||| | | | | | | | | | | | | | | | | | | | | |  
Db 1206 GTGTGAGATGAATGTGCTTTGAGCTTTAGTAAGATT 1244

RESULT 7  
US-09-510-332-84  
Sequence 84, Application US/09510332  
Publication No. US20030022278A1  
GENERAL INFORMATION:  
APPLICANT: Zuker, Charles S.  
APPLICANT: Adler, Jon Elliot  
APPLICANT: Ryba, Nick  
APPLICANT: Mueller, Ken  
APPLICANT: Hoon, Mark  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: T2R, a No. US20030022278A1el Family of Taste Receptors  
FILE REFERENCE: 02307E-0980100US  
CURRENT APPLICATION NUMBER: US/09/510,332  
CURRENT FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 09/393,634  
PRIOR FILING DATE: 1999-09-10  
NUMBER OF SEQ ID NOS: 172  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 84  
LENGTH: 3449  
TYPE: DNA  
ORGANISM: Rattus sp.  
FEATURE:  
OTHER INFORMATION: rat T2R04 (rGR04)  
US-09-510-332-84

Query Match 51.2%; Score 21; DB 9; Length 3449;  
Best Local Similarity 69.2%; Pred. No. 25;  
Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTGTGAGAGAGATGTGTCYCTAAGTGTGAGTGTGAGT 39  
||||| | | | | | | | | | | | | | | | | | | | | |  
Db 1206 GTGTGAGATGAATGTGCTTTGAGCTTTAGTAAGATT 1244

RESULT 8  
US-09-874-470-5  
Sequence 5, Application US/09874470  
Patent No. US20020071842A1  
GENERAL INFORMATION:  
APPLICANT: Gumperz, Jenny E  
APPLICANT: Brenner, Michael B  
APPLICANT: Behar, Samuel M  
TITLE OF INVENTION: Soluble CDI Compositions and Uses Thereof

```

; FILE REFERENCE: B00801/70212
; CURRENT APPLICATION NUMBER: US/09/874,470
; CURRENT FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,416
; PRIOR FILING DATE: 2000-06-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 10351
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-874-470-5

Query Match          50.7%; Score 20.8; DB 10; Length 10351;
Best Local Similarity 73.5%; Pred. No. 40;
Matches 25; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 8 AGAAGATGTCGYCTAAGTGCAGTGTCAGTCT 41
||||| ||| ||| ||| ||| |||
Db 5679 AGAAGATTCCTGCTCTCAGTCTCAGTCTTTGTCT 5712

RESULT 9
US-10-040-739-1149
; Sequence 1149, Application US/10040739
; Patent No. US20020173635A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John
; LaValle, Edward
; Racie, Lisa
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
; NUMBER OF SEQUENCES: 1519
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/040,739
; FILING DATE: 07-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/036,520
; FILING DATE: 03-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1149:
US-10-040-739-1149

Query Match          50.2%; Score 20.6; DB 9; Length 303;
Best Local Similarity 70.3%; Pred. No. 20;
```

```

Matches 26; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 2 TGTGAGAGAGAAATGTGTGCYCTAAGTGTTCAGTGTGAG 38
||||| ||| ||| ||| ||| ||| ||| |||
Db 127 TATGAGAGTGTGCGTGTGCTGTGAGTGTGTGTGGG 163

RESULT 10
US-09-949-842-6/c
; Sequence 6, Application US/09949842
; Patent No. US20020164692A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: PT047P1
; FILE REFERENCE: Immune System-Related Polynucleotides, Polypeptides, and Antibodie
; CURRENT APPLICATION NUMBER: US/09/949,842
; CURRENT FILING DATE: 2001-09-02
; PRIOR APPLICATION NUMBER: PCI/US01/07260
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/224,367
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/187,873
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 3980
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-949-842-6

Query Match          50.2%; Score 20.6; DB 9; Length 3980;
Best Local Similarity 70.3%; Pred. No. 38;
Matches 26; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 5 GAGAGAGAAATGTGTGCYCTAAGTGTTCAGTGTGAGTCT 41
|||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2595 GAAAGAAATGTGTGTGTGTGTGTGTGTGTGTGTGT 2559

RESULT 11
US-09-764-877-3691/c
; Sequence 3691, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3691
; LENGTH: 2331
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-877-3691

Query Match          49.8%; Score 20.4; DB 10; Length 2331;
Best Local Similarity 71.1%; Pred. No. 40;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 TGTGAGAGAGAAATGTGTGCYCTAAGTGTTCAGTGTGAGT 39
|||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1947 TCAAAAGGGAATTTTGTGTAAGTGTTCAGTGTAGT 1910

RESULT 12
US-09-942-325A-2
; Sequence 2, Application US/09942325A
; Patent No. US20020106794A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Iacovitti, Lorraine
; APPLICANT: Kessler, Mark
; TITLE OF INVENTION: The Human Tyrosine Hydroxylase Promoter
; FILE REFERENCE: IAC01.NP001
; CURRENT APPLICATION NUMBER: US/09/942,325A
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/228931
; PRIOR FILING DATE: 2000-02-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 10828
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-942-325A-2

Query Match          49.8%; Score 20.4; DB 10; Length 10828;
Best Local Similarity 67.5%; Pred. No. 58;
Matches 27; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 TGTGAGAGAGAAATGTGTCYCTAAGTGTGAGTGTGTCT 41
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 6360 TGTGTGTGCTTGTGTGCCCCCATGTGTCATGTGTGT 6399

RESULT 13
US-09-942-325A-1
; Sequence 1, Application US/09942325A
; Patent No. US20020106794A1
; GENERAL INFORMATION:
; APPLICANT: Iacovitti, Lorraine
; APPLICANT: Kessler, Mark
; TITLE OF INVENTION: The Human Tyrosine Hydroxylase Promoter
; FILE REFERENCE: IAC01.NP001
; CURRENT APPLICATION NUMBER: US/09/942,325A
; CURRENT FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/228931
; PRIOR FILING DATE: 2000-02-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 13329
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-942-325A-1

Query Match          49.8%; Score 20.4; DB 10; Length 13329;
Best Local Similarity 67.5%; Pred. No. 61;
Matches 27; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 TGTGAGAGAGAAATGTGTCYCTAAGTGTGAGTGTGTCT 41
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 6513 TGTGTGTGCTTGTGTGCCCCCATGTGTCATGTGTGT 6552

RESULT 14
US-10-094-989-3
; Sequence 3, Application US/10094989
; Patent No. US20020115179A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; FILE REFERENCE: CL001063DIV
; CURRENT APPLICATION NUMBER: US/10/094,989
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/754,250
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(111282)
; OTHER INFORMATION: n = A,T,C or G
US-10-094-989-3

Query Match          49.8%; Score 20.4; DB 12; Length 111282;
Best Local Similarity 71.1%; Pred. No. 1e+02;
Matches 27; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTCTGAGAGAGAAATGTGTGCYCTAAGTGTGAGTGTGAG 38
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 81928 GTCTCAGTGAGAAAGTGGGGGCTCAGTGAGAGGGGTGAG 81965

RESULT 15
US-09-864-761-7320
; Sequence 7320, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7320
; LENGTH: 513
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC015914.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.8
US-09-864-761-7320

Query Match          49.3%; Score 20.2; DB 10; Length 513;
Best Local Similarity 68.3%; Pred. No. 33;
Matches 28; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 GTGTGAGAGAAATGTGTGCYCYTAAGTGTCAAGTGTGAGTCT 41
   ||||| | | ||||| ||| ||||| || |
Db 462 GTGTGTGTGTGTGTGTGCGCTACAAGTCAGCTGGGTTT 502

Search completed: February 11, 2003, 09:10:30
Job time : 80.7214 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: February 10, 2003, 21:06:50 ; Search time 5781.15 Seconds  
(without alignments)  
8457.264 Million cell updates/sec

Title: US-09-942-310-2  
Perfect score: 1680  
Sequence: 1 gaattcaagaccagctgga.....catctctgctctctgtgg 1680

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

- 1: gb\_da.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pi.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_hgt\_hum.\*
- 31: em\_hgt\_inv.\*
- 32: em\_hgt\_other.\*
- 33: em\_hgt\_mus.\*
- 34: em\_hgt\_pln.\*
- 35: em\_hgt\_rod.\*
- 36: em\_hgt\_mam.\*
- 37: em\_hgt\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgtgo\_hum.\*
- 40: em\_hgtgo\_mus.\*
- 41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		DB	ID	Description	
		Match	Length				
1	1677.2	99.8	1680	6	AX394457	AX394457 Sequence	
2	1677.2	99.8	9432	6	AX394456	AX394456 Sequence	
3	1677.2	99.8	9432	6	HUMCYP2D6	M33388 Human cytoC	
4	1665.4	99.1	1669	6	AX207224	AX207224 Sequence	
5	1665.2	99.1	13677	9	HSCYP2D7B	X58468 Human CYP2D	
c	6	1220.4	72.6	114846	9	HS257120	AL021878 Human DNA
7	885.4	52.7	13278	9	HSCYP2D7A	X58467 Human CYP2D	
8	872.8	52.0	5503	9	HUMCYP2DG	M33189 Human debri	
9	866	51.5	17060	9	HUMCYP8P	M33387 Human debri	
10	756.4	45.0	5884	6	AX345458	AX345458 Sequence	
11	756.4	45.0	5884	6	AX348344	AX348344 Sequence	
c	12	537.4	32.0	5884	6	AX345459	AX345459 Sequence
c	13	537.4	32.0	5884	6	AX348345	AX348345 Sequence
c	14	284.8	17.0	78065	9	AC005355	AC005355 Homo sapi
c	15	280.6	16.7	191280	2	AC087389	AC087389 Homo sapi
c	16	280.6	16.7	193519	2	AC023123	AC023123 Homo sapi
17	280.6	16.7	216681	2	AC006534	AC006534 Homo sapi	
18	275.2	16.4	162813	2	AC021833	AC021833 Homo sapi	
c	19	273.6	16.3	125133	9	AC013274	AC013274 Homo sapi
c	20	271.2	16.1	129950	9	HS739H11	AL031289 Human DNA
21	268.8	16.0	132832	9	AC091934	AC091934 Homo sapi	
22	268.8	16.0	182557	2	AC091942	AC091942 Homo sapi	
c	23	268.2	16.0	163848	9	HS1056L3	AL031727 Human DNA
c	24	268	16.0	65385	2	AC123774	AC123774 Homo sapi
c	25	267.2	15.9	185311	9	AL355586	AL355586 Homo sapi
c	26	267	15.9	16475	9	AC104532	AC104532 Homo sapi
c	27	267	15.9	42572	2	AC004602	AC004602 Homo sapi
28	267	15.9	213294	2	AC099778	AC099778 Homo sapi	
29	266.8	15.9	117628	9	AC008891	AC008891 Homo sapi	
c	30	266.8	15.9	165593	2	AC026856	AC026856 Homo sapi
c	31	266.4	15.9	113872	9	HS57G9	295116 Human DNA s
c	32	265.6	15.8	147514	9	AC008521	AC008521 Homo sapi
c	33	265.6	15.8	219820	9	AC008738	AC008738 Homo sapi
c	34	265.4	15.8	137411	9	AC092066	AC092066 Homo sapi
c	35	265	15.8	147097	9	AC011317	AC011317 Homo sapi
c	36	265	15.8	172759	2	AC016142	AC016142 Homo sapi
37	265	15.8	174540	2	AC068801	AC068801 Homo sapi	
c	38	264.4	15.7	148876	9	AC011455	AC011455 Homo sapi
39	264.2	15.7	174474	2	AC067985	AC067985 Homo sapi	
c	40	264.2	15.7	175047	2	AC012521	AC012521 Homo sapi
c	41	264.2	15.7	181220	9	AC092060	AC092060 Homo sapi
c	42	264.2	15.7	217584	2	AC087749	AC087749 Homo sapi
c	43	263.6	15.7	97110	2	AC016815	AC016815 Homo sapi
c	44	263.6	15.7	208236	2	AC010536	AC010536 Homo sapi
45	263.4	15.7	168110	9	HS2427113	AL096701 Human DNA	

ALIGNMENTS

RESULT 1  
AX394457  
LOCUS AX394457  
DEFINITION Sequence 2 from Patent WO0218638.  
ACCESSION AX394457  
VERSION AX394457.1 GI:21065595  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1  
AUTHORS Risinger, C., Andersson, M.K., Lewander, T. and Ollasson, E.  
TITLE Detection of cyp2d6 polymorphisms  
JOURNAL Patent: WO 0218638-A 2 07-MAR-2002;



ORIGIN

Query Match 99.8%; Score 1677.2; DB 6; Length 9432;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1673; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

QY	1	GAATTCAAGACCGCTGGACAACTTGGGAAGACCSGGTCTCTACAAAAAATACAAAAT	60
DB	1	GAATTCAAGACCGCTGGACAACTTGGGAAGACCSGGTCTCTACAAAAAATACAAAAT	60
QY	61	AGCTGGATTGGGTGGGTGGCTATGCTTATATCCAGCACTTTGGGAGCCTGAGGTG	120
DB	61	AGCTGGATTGGGTGGGTGGCTATGCTTATATCCAGCACTTTGGGAGCCTGAGGTG	120
QY	121	GGTGGATCACCTGAAGTCAGGAGTTCGAAGACTAGCCTGGGCCAACATGGTGAACCCCTATC	180
DB	121	GGTGGATCACCTGAAGTCAGGAGTTCGAAGACTAGCCTGGGCCAACATGGTGAACCCCTATC	180
QY	181	TCTACTGAANAATAYAAAAAGCTAGACGTGGTGGCCACACACCTGTATCCAGCTACTTAG	240
DB	181	TCTACTGAANAATACAAAAGCTAGACGTGGTGGCCACACACCTGTATCCAGCTACTTAG	240
QY	241	GAGGCTGAGGCAGGAGATTTGCTTGAAGCCTAGAGGTGAAGTGTGTAGTGAGCCGAGATT	300
DB	241	GAGGCTGAGGCAGGAGATTTGCTTGAAGCCTAGAGGTGAAGTGTGTAGTGAGCCGAGATT	300
QY	301	GCATCATTTGCCAATGGAGGGAGCCACCAGCCTGGGCCAACAGAGAAATCTCCGCTCTC	360
DB	301	GCATCATTTGCCAATGGAGGGAGCCACCAGCCTGGGCCAACAGAGAAATCTCCGCTCTC	360
QY	361	CAAAAAAANAANAANAANAAGTATAGGCTGGGTGGTGCCTGTAGTCCCAAGCTACTT	420
DB	361	CAAAAAAANAANAANAANAAGTATAGGCTGGGTGGTGCCTGTAGTCCCAAGCTACTT	420
QY	421	GGGAGCAGGGGTCCACTTGTATGTCGAGACTGCAGTGCAGCCATGATCCTGCCACTGCAC	480
DB	421	GGGAGCAGGGGTCCACTTGTATGTCGAGACTGCAGTGCAGCCATGATCCTGCCACTGCAC	480
QY	481	TCCGGCTGGGCACACAGAGTGAGACCTGTCTAAAGAAAAAANAANAAGCAACATATC	540
DB	481	TCCGGCTGGGCACACAGAGTGAGACCTGTCTAAAGAAAAAANAANAAGCAACATATC	540
QY	541	CTGAACAAAGGATCCTCCATACGTTTCCACACAGATTTCTAATCAGAAACATGGAGCCA	600
DB	541	CTGAACAAAGGATCCTCCATACGTTTCCACACAGATTTCTAATCAGAAACATGGAGCCA	600
QY	601	GAAAGCAGTGGAGGAGGACRACCTCAGCAGCCCGGGAGGATGTTGTCACAGGCTGGGG	660
DB	601	GAAAGCAGTGGAGGAGGAGCACCCTCAGCAGCCCGGGAGGATGTTGTCACAGGCTGGGG	660
QY	661	CAAGGCCCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTGGCAACAAGAAGC	720
DB	661	CAAGGCCCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTGGCAACAAGAAGC	720
QY	721	CATAGCCCGCCAGACCCAGGAATGTGGGTGGGCTGGGAGCAGCCTCTGGACAGAGT	780
DB	721	CATAGCCCGCCAGACCCAGGAATGTGGGTGGGCTGGGAGCAGCCTCTGGACAGAGT	780
QY	781	GGTCCCATCCAGGAACCTCCGGCATGGCTGGGAAGTGGGCTACTTGGTGGCGGCTGT	840
DB	781	GGTCCCATCCAGGAACCTCCGGCATGGCTGGGAAGTGGGCTACTTGGTGGCGGCTGT	840
QY	841	ATGTGTGTGACTGTGTCTGTGAGAGAGATGTGTGCYCTAAGTGTCTAGTGTGAGTCT	900
DB	841	ATGTGTGTGACTGTGTGTGTGAGAGAGATGTGTGCYCTAAGTGTCTAGTGTGAGTCT	900
QY	901	GTGTATGTGAATATTTCTTTGTGGGTGATTTTCTGCTGTGTGTGTGTGTGTGTGTGT	960
DB	901	GTGTATGTGAATATTTCTTTGTGGGTGATTTTCTGCTGTGTGTGTGTGTGTGTGTGT	960
QY	961	CAAGTGTGAACAGTGCACAACTGTCTGGAGTGGACAGAGATCTGTCCACCATCAGT	1020
DB	961	CAAGTGTGAACAGTGCACAACTGTCTGGAGTGGACAGAGATCTGTCCACCATCAGT	1020

QY	1021	GTCTGCATAGCGTCTGTGTCATGTCAAGAGTGCAGAGTGAAGTGAAGGGACCGCCCATG	1080
DB	1021	GTCTGCATAGCGTCTGTGTCATGTCAAGAGTGCAGAGTGAAGTGAAGGGACCGCCCATG	1080
QY	1081	ATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGTAAAGTGCAGCTGACAGATGAAGGTG	1140
DB	1081	ATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGTAAAGTGCAGCTGACAGATGAAGGTG	1140
QY	1141	CTCAAGTCTACTCTGGAGTGGCAGGTGGGGTAGGGAAGGCAAGGCCATGTTCTTGA	1200
DB	1141	CTCAAGTCTACTCTGGAGTGGCAGGTGGGGTAGGGAAGGCAAGGCCATGTTCTTGA	1200
QY	1201	GGAGGGTTTGTGACTACATTAGGGTGTATGAGCCTAGCTGGGAGGTGGATGGCCRGSTCC	1260
DB	1201	GGAGGGTTTGTGACTACATTAGGGTGTATGAGCCTAGCTGGGAGGTGGATGGCCRGSTCC	1260
QY	1261	ACTGAACCCCTGGTTATCCAGAGGCTTTGAGGCTTCAGGAGCTTGGAGTGGGAGAG	1320
DB	1261	ACTGAACCCCTGGTTATCCAGAGGCTTTGAGGCTTCAGGAGCTTGGAGTGGGAGAG	1320
QY	1321	GGGTGACTTCTCCGACAGGCGCCCTCCACCGGCTTACCTGGGTAAAGGCCCTGGAGCAG	1380
DB	1321	GGGTGACTTCTCCGACAGGCGCCCTCCACCGGCTTACCTGGGTAAAGGCCCTGGAGCAG	1380
QY	1381	GAAAGCAGGGGCAAGAACCTCTGGAGCAGCCCATACCGCCCTGCCCTGACTCTGCCACTG	1440
DB	1381	GAAAGCAGGGGCAAGAACCTCTGGAGCAGCCCATACCGCCCTGCCCTGACTCTGCCACTG	1440
QY	1441	GCAGCAGCTCAACACAGCAGGTTCACACAGAGGCAAGGCCATCATCAGCTCC	1500
DB	1441	GCAGCAGCTCAACACAGCAGGTTCACACAGAGGCAAGGCCATCATCAGCTCC	1500
QY	1501	CTTTATAAGGGAAGGTCACGCGCTCGGTGCTGAGAGTGTCTGCCCTGCTCTGTG	1560
DB	1501	CTTTATAAGGGAAGGTCACGCGCTCGGTGCTGAGAGTGTCTGCCCTGCTCTGTG	1560
QY	1561	CCTGTGGGTGGGGTGGCCAGGTGTCTCAGAGAGGCCATTTGGTAGTGAGGCAGGTA	1620
DB	1561	CCTGTGGGTGGGGTGGCCAGGTGTCTCAGAGAGGCCATTTGGTAGTGAGGCAGGTA	1620
QY	1621	TGGGGCTAGAGCAGCTGTGGCCCTGGCCGTGTATGTAGTGCCCATCTTCTGCTGCTGGTG	1680
DB	1621	TGGGGCTAGAGCAGCTGTGGCCCTGGCCGTGTATGTAGTGCCCATCTTCTGCTGCTGGTG	1680

RESULT 3  
LOCUS HUMCYP2D6  
DEFINITION Human cytochrome P450 IID6 (CYP2D6) gene, complete cds.  
ACCESSION M33388  
VERSION 1  
KEYWORDS cytochrome P450; cytochrome P450 IID6.  
SOURCE Human DNA, clone lambda2D-18/2.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 9432)  
AUTHORS Kimura S., Umemo M., Skoda R.C., Meyer U.A. and Gonzalez F.J.  
TITLE The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene

JOURNAL Am. J. Hum. Genet. 45 (6), 889-904 (1989)  
MEDLINE 90072069  
PUBMED 2574001

COMMENT Draft entry and computer-readable sequence for [Am. J. Hum. Genet. 45, 889-904 (1989)] kindly submitted by S.Kimura, 29-MAR-1990.

FEATURES  
Location/Qualifiers  
1..9432  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="22q13.1"

HUMCYP2D6 9432 bp DNA linear PRI 22-NOV-1994  
Human cytochrome P450 IID6 (CYP2D6) gene, complete cds.

M33388.1 GI:181303  
cytochrome P450; cytochrome P450 IID6.  
Human DNA, clone lambda2D-18/2.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 9432)  
Kimura S., Umemo M., Skoda R.C., Meyer U.A. and Gonzalez F.J.  
The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene  
Am. J. Hum. Genet. 45 (6), 889-904 (1989)  
90072069  
2574001

Draft entry and computer-readable sequence for [Am. J. Hum. Genet. 45, 889-904 (1989)] kindly submitted by S.Kimura, 29-MAR-1990.  
Location/Qualifiers  
1..9432  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="22q13.1"



QY 661 CAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGACAGCCCTGTGTCACAAAGAAGC 720  
Db 661 CAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGACAGCCCTGTGTCACAAAGAAGC 720  
QY 721 CATAGCCGGCCAGAGCCAGGAAATGTGGCTGGGAGCAGCCTCTGGACAGAGT 780  
Db 721 CATAGCCGGCCAGAGCCAGGAAATGTGGCTGGGAGCAGCCTCTGGACAGAGT 780  
QY 781 GGTCCTCATCCAGAAACCTCCGGCATGGCTGGGAAGTGGGTAATCTGGTCCCGGCTGT 840  
Db 781 GGTCCTCATCCAGAAACCTCCGGCATGGCTGGGAAGTGGGTAATCTGGTCCCGGCTGT 840  
QY 841 ATGTGTGTGACTGTGTGTGTGAGAGAAATGTGTGCTGAGTGTGAGTGT 900  
Db 841 ATGTGTGTGACTGTGTGTGTGAGAGAAATGTGTGCTGAGTGTGAGTGT 900  
QY 901 GTGTATGTGTGAATATATGTCTTTGTGTGGGTGATTTTCTGCTGTGTAACTGGTCCCTG 960  
Db 901 GTGTATGTGTGAATATATGTCTTTGTGTGGGTGATTTTCTGCTGTGTAACTGGTCCCTG 960  
QY 961 CAAGTGTGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGT 1020  
Db 961 CAAGTGTGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGT 1020  
QY 1021 GTGTGCATAGCGTCTGTGCATGTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1080  
Db 1021 GTGTGCATAGCGTCTGTGCATGTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1080  
QY 1081 ATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGTGAAGTGAAGTGAAGTGAAGTGAAGT 1140  
Db 1081 ATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGTGAAGTGAAGTGAAGTGAAGTGAAGT 1140  
QY 1141 CTGAAGTCACTCTGGAGTGGGAGTGGGGTAGGGAAGGCAAGGCCATGTCTTCTGGA 1200  
Db 1141 CTGAAGTCACTCTGGAGTGGGAGTGGGGTAGGGAAGGCAAGGCCATGTCTTCTGGA 1200  
QY 1201 GGAGGGGTGTGTACTATAGGTGTATGAGCCTAGTCTGGGAGTGGATGGCCRGCTC 1260  
Db 1201 GGAGGGGTGTGTACTATAGGTGTATGAGCCTAGTCTGGGAGTGGATGGCCRGCTC 1260  
QY 1261 ACTGAACCCCTGGTTATCCAGAGGCTTTGCAGGCTTTCAGGAGTTCAGGAGTGGGAGAG 1320  
Db 1261 ACTGAACCCCTGGTTATCCAGAGGCTTTGCAGGCTTTCAGGAGTTCAGGAGTGGGAGAG 1320  
QY 1321 GGGGTGACTTCTCCGACAGGCCCTCCACCGGCTACCCCTGGGTAAAGGCCCTGGAGCAG 1380  
Db 1321 GGGGTGACTTCTCCGACAGGCCCTCCACCGGCTACCCCTGGGTAAAGGCCCTGGAGCAG 1380  
QY 1381 GAAGCAGGGCCAGAACCTCTGAGCAGGCCCATACCGGCCCTGGCTGACTCTGCCACTG 1440  
Db 1381 GAAGCAGGGCCAGAACCTCTGAGCAGGCCCATACCGGCCCTGGCTGACTCTGCCACTG 1440  
QY 1441 GCAGCAGTCAACACAGCAGTTCACCTACAGCAGAGGCCAAGGCCATCATCAGCTCC 1500  
Db 1441 GCAGCAGTCAACACAGCAGTTCACCTACAGCAGAGGCCAAGGCCATCATCAGCTCC 1500  
QY 1501 CTTTATAAGGGAAGGTCACGCGCTCGGTGTGTGTGAGTGTCTCTGCTGCTCTGTG 1560  
Db 1501 CTTTATAAGGGAAGGTCACGCGCTCGGTGTGTGTGAGTGTCTCTGCTGCTCTGTG 1560  
QY 1561 CTTGTGGGTGGGGTGCAGGTGTGTCCAGAGGCCCATTTGGTAGTGAGGAGGTA 1620  
Db 1561 CTTGTGGGTGGGGTGCAGGTGTGTCCAGAGGCCCATTTGGTAGTGAGGAGGTA 1620  
QY 1621 TGGGGCTAGAAGCAGTGGTCCCTGGCCGTGATAGTGGCCATCTTCTCTCTCTGGTGG 1680  
Db 1621 TGGGGCTAGAAGCAGTGGTCCCTGGCCGTGATAGTGGCCATCTTCTCTCTCTGGTGG 1680

RESULT 4  
AX207224  
LOCUS AX207224 1669 bp DNA linear PAT 30-AUG-2001  
DEFINITION Sequence 1 from Patent WO0155432.

ACCESSION AX207224  
VERSION AX207224.1 GI:15394976  
KEYWORDS  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE artificial sequences.  
1 (bases 1 to 1669)  
AUTHORS Raimundo,S. and Zanger,U.  
TITLE Polymorphisms in the human cyp2d6 gene promoter region and their use in diagnostic and therapeutic applications  
JOURNAL Patent: WO 015432-A 1 02-AUG-2001;  
Epidaurus Biotechnologie AG (DE)  
FEATURES  
source Location/Qualifiers  
1..1669  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="artificial sequence"  
BASE COUNT 413 a 376 c 534 g 338 t 8 others  
ORIGIN  
Query Match 99.1%; Score 1665.4; DB 6; Length 1669;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1666; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAATTCAAGACCACGCTGGACAACCTTGGAAAGAACCCGGTCTCTACAAAAAATACAAAATT 60  
Db 1 GAATTCAAGACCACGCTGGACAACCTTGGAAAGAACCCGGTCTCTACAAAAAATACAAAATT 60  
QY 61 AGCTGGGATTGGGTGGCGTGCCTATTAATCCACGACACTTTGGGAGCCTGAGGTG 120  
Db 61 AGCTGGGATTGGGTGGCGTGCCTATTAATCCACGACACTTTGGGAGCCTGAGGTG 120  
QY 121 GGTGGATCACCTCAAGTTCAGGAGTTCAAAGACTAGCCCTGGCCAAACATGGTGAACCCCTATC 180  
Db 121 GGTGGATCACCTCAAGTTCAGGAGTTCAAAGACTAGCCCTGGCCAAACATGGTGAACCCCTATC 180  
QY 181 TCTACTGAAAATAYAAAAAGCTAGACGTGTGGCACACACCTGTAAATCCAGCTACTTAG 240  
Db 181 TCTACTGAAAATAYAAAAAGCTAGACGTGTGGCACACACCTGTAAATCCAGCTACTTAG 240  
QY 241 GAGGCTGAGCAGAGAAATTCCTGAAGCCTAGAGGTGAAGGTGTGATGAGCGGAGATT 300  
Db 241 GAGGCTGAGCAGAGAAATTCCTGAAGCCTAGAGGTGAAGGTGTGATGAGCGGAGATT 300  
QY 301 GCATCATTTGCACAAATGGAGGGAGCCACACGCTGGGCCAAAGAGGAAATCTCCGCTCTC 360  
Db 301 GCATCATTTGCACAAATGGAGGGAGCCACACGCTGGGCCAAAGAGGAAATCTCCGCTCTC 360  
QY 361 CAAAAAATAAAAAAATAAAAAAAGATTAGGCTGGGTGGTGCCTGTAGTCCCAGCTACTT 420  
Db 361 CAAAAAATAAAAAAATAAAAAAAGATTAGGCTGGGTGGTGCCTGTAGTCCCAGCTACTT 420  
QY 421 GGAGGCGAGGGGTCCACTTGTATGTCCAGACTGCAGTGGAGCCATGATCTCCGCTGCAC 480  
Db 421 GGAGGCGAGGGGTCCACTTGTATGTCCAGACTGCAGTGGAGCCATGATCTCCGCTGCAC 480  
QY 481 TCCGGCTTGGCCAAACAGAGTGAGACCTGTCTAAAGAAAAAATAAAGCAACATATC 540  
Db 481 TCCGGCTTGGCCAAACAGAGTGAGACCTGTCTAAAGAAAAAATAAAGCAACATATC 540  
QY 541 CTGAACAAAGGATCTCCATAAGCTTCCCACAGATTTCTAATCAGAAACATGAGGCCA 600  
Db 541 CTGAACAAAGGATCTCCATAAGCTTCCCACAGATTTCTAATCAGAAACATGAGGCCA 600  
QY 601 GAAAGCAGTGGAGGAGGACRACCTCAGGCAGCCCGGGAGGATTTGTCTACAGGCTGGGG 660  
Db 601 GAAAGCAGTGGAGGAGGACRACCTCAGGCAGCCCGGGAGGATTTGTCTACAGGCTGGGG 660  
QY 661 CAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTGTCACAGGCTGGGG 720  
Db 661 CAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTGTCACAGGCTGGGG 720  
QY 721 CATAGCCGGCCAGAGCCAGGAAATGTGGCTGGGAGTGGGTAATCTGGTCCCGGCTGT 780



Intron	4923..5376			
	/gene="CYP2D7BP"			
	/number=7			
exon	5377..5518			
	/gene="CYP2D7BP"			
	/number=8			
Intron	5519..5616			
	/gene="CYP2D7BP"			
	/number=8			
exon	5617..5868			
	/gene="CYP2D7BP"			
	/number=9			
repeat_region	8267..8306			
	/note="ACCCCTCCCC"			
BASE COUNT	3066 a	3775 c	4107 g	2729 t
ORIGIN				
Query Match	99.1%;	Score 1665.2;	DB 9;	Length 13677;
Best Local Similarity	99.5%;	Pred. No. 0;		
Matches 1673;	Conservative 7;	Mismatches 0;	Indels 2;	Gaps 1;
QY	1	GAATTCAGACACAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAAATACAAAAAT	60	
Db	1	GAATTCAGACACAGCCTGGACAACTTGGAGAACCCGGTCTCTACAAAAAATACAAAAAT	60	
QY	61	AGCTGGGATTGGGTGCGGTGGCTCATGCCCTATAATCCCAGCACTTTGGAGCCTGAGGTG	120	
Db	61	AGCTGGGATTGGGTGCGGTGGCTCATGCCCTATAATCCCAGCACTTTGGGAGCCTGAGGTG	120	
QY	121	GGTGGATCACCTGAAGTCAGGAGTTCGAAGACTAGCCTGGCCCAACATGGTGAACCCCTATC	180	
Db	121	GGTGGATCACCTGAAGTCAGGAGTTCGAAGACTAGCCTGGCCCAACATGGTGAACCCCTATC	180	
QY	181	TCTACTGAAAATAYAAAAAGCTAGAGCTGGTGGCACACACCTGTAATCCAGCTACTTAG	240	
Db	181	TCTACTGAAAATATAAAAAGCTAGAGCTGGTGGCACACACCTGTAATCCAGCTACTTAG	240	
QY	241	GAGGCTGAGGCAGGAGAAATGCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCCGAGATT	300	
Db	241	GAGGCTGAGGCAGGAGAAATGCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCCGAGATT	300	
QY	301	GCATCATTTGCACAAATGAGGGGAGCCACCAGCCTGGGCCAACAGAGAAATCTCCGCTCTC	360	
Db	301	GCATCATTTGCACAAATGAGGGGAGCCACCAGCCTGGGCCAACAGAGAAATCTCCGCTCTC	360	
QY	361	C--AAAAAATAAAAAAAGATATAGGCTGGGTGGTGCCTGTAGTCCCACTAC	418	
Db	361	CAAAAAAATAAAAAAAGATATAGGCTGGGTGGTGCCTGTAGTCCCACTAC	420	
QY	419	TTGGGAGGCAGGGGTCCACTTGATGTCGAGACTGCAGTGAGCCATGATCCTGCCACTGC	478	
Db	421	TTGGGAGGCAGGGGTCCACTTGATGTCGAGACTGCAGTGAGCCATGATCCTGCCACTGC	480	
QY	479	ACTCCGGCCTGGGCACACAGAGTGAGACCTGTCTAAAGAAAAAATAAGCAACATA	538	
Db	481	ACTCCGGCCTGGGCACACAGAGTGAGACCTGTCTAAAGAAAAAATAAGCAACATA	540	
QY	539	TCCTGAACAAAGGATCCTCCATAACGTTCCCAACAGATTTCCTAATCAGAAACATGGAGGC	598	
Db	541	TCCTGAACAAAGGATCCTCCATAACGTTCCCAACAGATTTCCTAATCAGAAACATGGAGGC	600	
QY	599	CAGAAAGCAGTGAGGAGGACRACCCCTCAGGCAGCCGGGAGGATGTTGTACAGGCTGG	658	
Db	601	CAGAAAGCAGTGAGGAGGACRACCCCTCAGGCAGCCGGGAGGATGTTGTACAGGCTGG	660	
QY	659	GGCAAGGCCCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTTGCAAAACAGAA	718	
Db	661	GGCAAGGCCCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTTGCAAAACAGAA	720	
QY	719	GCCATAGCCCGGCCAGAGCCAGGAATGTGGCTGGGCTGGGAGCAGCCTCTGGACAGGA	778	
Db	721	GCCATAGCCCGGCCAGAGCCAGGAATGTGGCTGGGCTGGGAGCAGCCTCTGGACAGGA	780	

RESULT 6  
HS257I20/c  
LOCUS  
DEFINITION  
22q13.1-13.2, complete sequence.  
ACCESSION  
AL021878  
VERSION  
AL021878.2  
KEYWORDS  
HTG.

HS257I20  
Human DNA sequence from clone RP1-257I20 on chromosome  
114846 bp DNA linear  
PRI 22-NOV-2001

QY	779	GTGTCCTCCATCCAGAAACCTCCGGCATGGCTGGGAAGTGGGTACTTGGTGCCTGCTCT	838	
Db	781	GTGTCCTCCATCCAGAAACCTCCGGCATGGCTGGGAAGTGGGTACTTGGTGCCTGCTCT	840	
QY	839	GTATGTGTGTGACTGGTGTGTGTGAGAGAGAAATGTGTCYCTAAGTGTCACTGTCACT	898	
Db	841	GTATGTGTGTGACTGGTGTGTGTGAGAGAGAAATGTGTCYCTAAGTGTCACTGTCACT	900	
QY	899	CTGTGTATGTGTAATATTGCTTTTGTGGGTGATTTTCTGCTGTGTAAATGCTGTCTCC	958	
Db	901	CTGTGTATGTGTAATATTGCTTTTGTGGGTGATTTTCTGCTGTGTAAATGCTGTCTCC	960	
QY	959	TGCAAGTGTGAACAAGTGGACAAAGTCTCTGGGAGTGGACAAGAGATCTGTGCACCATCAG	1018	
Db	961	TGCAAGTGTGAACAAGTGGACAAAGTCTCTGGGAGTGGACAAGAGATCTGTGCACCATCAG	1020	
QY	1019	GTGTGTGCATAGCGTCTGTCCATGTCACAGTGCACAGTGAAGTGAAGGACACAGGCCCA	1078	
Db	1021	GTGTGTGCATAGCGTCTGTCCATGTCACAGTGAAGTGAAGGACACAGGCCCA	1080	
QY	1079	TGATGCCACTCATCATCAGGAGCTCTAAGGCCCAAGTAAAGTCCAGTGCACAGATAAGGG	1138	
Db	1081	TGATGCCACTCATCATCAGGAGCTCTAAGGCCCAAGTAAAGTCCAGTGCACAGATAAGGG	1140	
QY	1139	TGCTGAAGTCACTCTGAGTGGGAGGTGGGGTAGGAAAGGCAAGGCCCATGTTCTG	1198	
Db	1141	TGCTGAAGTCACTCTGAGTGGGAGGTGGGGTAGGAAAGGCAAGGCCCATGTTCTG	1200	
QY	1199	GAGGAGGGTGTGACTACATTAAGGTGTATGAGCCTAGCTGGGAGTGGATGCCRGGT	1258	
Db	1201	GAGGAGGGTGTGACTACATTAAGGTGTATGAGCCTAGCTGGGAGTGGATGCCRGGT	1260	
QY	1259	CACTCAAAACCTTGTATCCCAAGAGGCTTTGCAGGCTTCAGAGCTTGGAGTGGGGAG	1318	
Db	1261	CACTCAAAACCTTGTATCCCAAGAGGCTTTGCAGGCTTCAGAGCTTGGAGTGGGGAG	1320	
QY	1319	AGGGGTGACTTCTCCGACACAGCCCTCCACGGCTACCCCTGGGTAAAGGCCCTGAGC	1378	
Db	1321	AGGGGTGACTTCTCCGACACAGCCCTCCACGGCTACCCCTGGGTAAAGGCCCTGAGC	1380	
QY	1379	AGGAACAGGGGCAAGAACTCTGGAGACAGCCCATACCCGCCCTGGCCTGACTCTGCCAC	1438	
Db	1381	AGGAACAGGGGCAAGAACTCTGGAGACAGCCCATACCCGCCCTGGCCTGACTCTGCCAC	1440	
QY	1439	TGCAACACAGTCAACACAGCAGTTTCACTCACAGCAGAGGGCAAGGCCATCATCAGCT	1498	
Db	1441	TGCAACACAGTCAACACAGCAGTTTCACTCACAGCAGAGGGCAAGGCCATCATCAGCT	1500	
QY	1499	CCCTTTATAGGGAAGGGTCACGCGCTCGGTGTGTGTGAGAGTGTCTGCTGCTCTCTG	1558	
Db	1501	CCCTTTATAGGGAAGGGTCACGCGCTCGGTGTGTGTGAGAGTGTCTGCTGCTCTCTG	1560	
QY	1559	TGCTCGTGGGTGGGGTCCAGAGTGTGTCCAGAGAGCCCATTTGGTAGTCAGGAGG	1618	
Db	1561	TGCTCGTGGGTGGGGTCCAGAGTGTGTCCAGAGAGCCCATTTGGTAGTCAGGAGG	1620	
QY	1619	TATGGGCTAGAGACACTGTGTCCTGCGCCCTGGCCGTGATAGTGGCCATCTTCTGCTCTGT	1678	
Db	1621	TATGGGCTAGAGACACTGTGTCCTGCGCCCTGGCCGTGATAGTGGCCATCTTCTGCTCTGT	1680	
QY	1679	GG 1680		
Db	1681	GG 1682		



```
SOURCE
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 114846)
AUTHORS      Bridgeman,A.
TITLE        Direct Submission
JOURNAL      Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
COMMENT
On Nov 25, 2001 this sequence version replaced gi:3204432.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep
This sequence
was finished as follows unless otherwise noted: all regions were
either double-stranded or sequenced with an alternate chemistry or
covered by high quality data (i.e., phred quality >= 30); an
attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by at least one
plasmid subclone or more than one M13 subclone; and the assembly
was confirmed by restriction digest. This sequence was generated
from part of bacterial clone contigs of human chromosome 22,
constructed by the Sanger Centre Chromosome 22 Mapping Group.
Further information can be found at
http://www.sanger.ac.uk/HGP/Chr22
RPI-257120 is from the library RPCI-1 constructed by the group of
Pietter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: PCPAC2
This sequence is the entire insert of clone RPI-257120 The true
right end of clone RPI-18601 is at 20171 in this sequence.
FEATURES
source
1..114846
/organism="Homo sapiens"
/db_xref="RZPD:RPI-18601"
/db_xref="taxon:9606"
/chromosomes="22"
/map="q13.1-13.2"
/clone="RPI-257120"
/clone_lib="RPCI-1"
misc_feature
1093..1682
/note="match: GSS: Em:AQ887153"
1097..1600
/note="match: GSS: Em:AQ563517"
complement(3894..4138)
/note="match: GSS: Em:AQ140240"
4148..4647
/note="match: GSS: Em:AQ684484"
7162..7366
/note="match: STS: Em:HS324WC5"
7357..7396
/note="20 copies 2 mer ac 82% conserved"
7369..7551
/note="match: STS: Em:HS324WC5"
7972..9283
/note="CpG island"
/evidence=not_experimental
misc_feature
9981..10378
/note="match: GSS: Em:AQ124532"
13208..14275
/note="MER11C repeat: matches 1..1057 of consensus"
16086..16507
/note="L1MD repeat: matches 1552..1964 of consensus"
17055..17127
/note="L1MD repeat: matches 1482..1552 of consensus"
```

```
repeat_region
17434..17688
/note="L1MD repeat: matches 1222..1482 of consensus"
19152..19153
/note="clone RPI-18601
aa in this entry
deletion"
/replace="aaaaa"
19896..19939
/note="L1P repeat: matches 2416..2459 of consensus"
20246..20284
/note="L1P repeat: matches 2378..2416 of consensus"
20594..20660
/note="L1P repeat: matches 2311..2378 of consensus"
20965..20982
/note="L1P repeat: matches 2294..2311 of consensus"
21270..22130
/note="L1P repeat: matches 1445..2294 of consensus"
22425..22499
/note="L1P repeat: matches 1370..1445 of consensus"
22502..22994
/note="L1P repeat: matches 1..510 of consensus"
23020..25379
/note="BaEV-int repeat: matches 4243..6625 of consensus"
25398..25662
/note="match: STS: Em:G07321"
25597..26083
/note="match: GSS: Em:B13982"
25611..25853
/note="match: GSS: Em:B14069"
complement(26269..26446)
/note="match: GSS: Em:AF046780"
26890..26995
/note="match: GSS: Em:AZ083430"
27147..28076
/note="HERV1 repeat: matches 4455..5408 of consensus"
28386..30853
/note="HERV1 repeat: matches 2098..4455 of consensus"
30843..30912
/note="HERV1 repeat: matches 2098..2163 of consensus"
30902..30971
/note="HERV1 repeat: matches 2098..2163 of consensus"
30961..31556
/note="HERV1 repeat: matches 1561..2163 of consensus"
31873..33456
/note="HERV1 repeat: matches 9..1561 of consensus"
33464..33959
/note="L1R10B repeat: matches 1..510 of consensus"
33960..34110
/note="L1PBA repeat: matches 1226..1378 of consensus"
34415..35367
/note="L1PBA repeat: matches 656..1226 of consensus"
37163..37323
/note="Hariequin repeat: matches 912..1086 of consensus"
37324..37437
/note="57 copies 2 mer ga 79% conserved"
complement(join(42379..42591,42916..43112))
/note="match: STS: Em:G27508"
complement(join(42491..42603,42916..42945))
/note="match: STS: Em:G43129"
43992..44915
/note="PTR5 repeat: matches 40..93 of consensus"
44916..45564
/note="L1R12 repeat: matches 5..671 of consensus"
48350..49467
/note="CpG island"
/evidence=not_experimental
54417..54450
/note="17 copies 2 mer ca 100% conserved"
54578..55083
/note="MER77 repeat: matches 30..587 of consensus"
58051..59086
/note="CpG island"
/evidence=not_experimental
```









Db	361	TGGGGGTAGGGAAGGCGCAGCCGATGTTCTGGAGGAGGGGTTGTGACTACATTAGGGTG	420		exon	/product="debrisoquine 4-hydroxylase"
						<1392..1568
QY	1227	TATGAGCCTAGCTGGAGTGGTCCACGTCCACTGAACCCCTGCTTATCCAGAGG	1286		exon	/gene="CYP2D8p"
						/number=1
Db	421	TATGAGCCTAGCTGGAGTGGTCCGCGGTCCACTGAACCCCTGCTTATCCAGAGG	480			3189..3360
						/gene="CYP2D8p"
						/number=2
QY	1287	CTTTGCAGGCTTCAGGAGCTTTGAGTGGGGAGAGGGGCTGACTTCTCCGACAGGCCCT	1346		exon	/pseudo
						3907..4059
Db	481	CTTTGCAGGCTTCAGGAGCTTTGAGTGGGGAGAGGGGCTGACTTCTCCGACAGGCCCT	540			/gene="CYP2D8p"
						/number=3
QY	1347	CCACGGCCTACCTCGGTAGGCGCTGGACGAGGAGGCGCAAGAACCTCTGGAGC	1406		exon	/pseudo
						4148..4310
Db	541	CCACGGCCTACCTCGGTAGGCGCTGGACGAGGAGGCGCAAGAACCTCTGGAGC	600			/gene="CYP2D8p"
						/number=4
QY	1407	AGCCCATACCGCCCTGGCTGACTCTGCCACTGGCAGCACAGTCAACACAGCAGGTTCA	1466		exon	/pseudo
						4758..4934
Db	601	AGCCCATACCGCCCTGGCTGACTCTGCCACTGGCAGCACAGTCAACACAGCAGGTTCA	660			/gene="CYP2D8p"
						/number=5
QY	1467	CTCACAGCAGAGGCGCAAGGCCATCATCAGTCCCTTTATAAGGGAAGGTCACGGCGTC	1526		exon	/pseudo
						5121..5262
Db	661	CTCACAGCAGAGGCGCAAGGCCATCATCAGTCCCTTTATAAGGGAAGGTCACGGCGTC	720			/gene="CYP2D8p"
						/number=6
QY	1527	GGTGTCTGAGAGTGCTCCTGCCTGGTCTCTGTGCCCTGGTGGGGTGGCGAGGTGT	1586		exon	/pseudo
						5467..5651
Db	721	GGTGTCTGAGAGTGCTCCTGCCTGGTCTCTGTGCCCTGGTGGGGTGGCGAGGTGT	780			/gene="CYP2D8p"
						/number=7
QY	1587	GTCACAGAGGCCATTTGTTAGTGAGGCAGGTATGGGCTAGAACACACGTGGTGGCCCTG	1646		exon	/pseudo
						6101..6242
Db	781	GTCACAGAGGCCATTTGTTAGTGAGGCAGGTATGGGCTAGAACACACGTGGTGGCCCTG	840			/gene="CYP2D8p"
						/number=8
QY	1647	GCGGTGATAGTGCCCAFTTCCTGCTCCTGGTGG 1680			exon	/pseudo
						6339..>6516
Db	841	GCGGTGATAGTGCCCAFTTCCTGCTCCTGGTGG 874				/gene="CYP2D8p"
						/number=9
RESULT 9						
HUMCYPBP						
LOCUS						
DEFINITION	HUMCYP8P	17060 bp	DNA	linear	PRI 09-NOV-1994	
	Human debrisoquine 4-hydroxylase (CYP2D8p) and (CYP2D7) pseudogenes					
	complete sequences.					
ACCESSION	M33387					
VERSION	M33387.1	GI:181320				
KEYWORDS	debrisoquine 4-hydroxylase.					
SOURCE	Human DNA, clones lambda-2D-A and lambda-2D-B.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 17060)					
AUTHORS	Kimura,S., Umeno,M., Skoda,R.C., Meyer,U.A. and Gonzalez,F.J.					
TITLE	The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and					
	identification of the polymorphic CYP2D6 gene, a related gene, and					
	a pseudogene					
JOURNAL	Am. J. Hum. Genet. 45 (6), 889-904 (1989)					
MEDLINE	90072069					
PUBMED	2574001					
COMMENT	Draft entry and computer-readable sequence for [1] kindly submitted					
	by S.Kimura, 29-MAR-1990.					
FEATURES	Location/Qualifiers					
source	1..17060					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
TATA_signal	1276..1282					
prim_transcript	1304..6570					
	/note="CYP2D8p mRNA and introns"					
gene	Join(1392..1568,3189..3360,3907..4059,4148..4310,					
	4758..4934,5121..5262,5467..5651,6101..6242,6339..6516)					
	/gene="CYP2D8p"					
CDS	Join(1392..1568,3189..3360,3907..4059,4148..4310,					
	4758..4934,5121..5262,5467..5651,6101..6242,6339..6516)					
	/gene="CYP2D8p"					
	/pseudo					
	/codon_start=1					

```
/gene="CYP2D7"  
/number-7  
/pseudo  
15080..15221  
/gene="CYP2D7"  
/number-8  
/pseudo  
15320..>15496  
/gene="CYP2D7"  
/number-9  
/number-10  
BASE COUNT      3516 a  4595 c  5034 g  3915 t  
ORIGIN  
Query Match      51.5%; Score 866; DB 9; Length 17060;  
Best Local Similarity 96.3%; Pred. No. 4.7e-239;  
Matches 894; Conservative 3; Mismatches 28; Indels 3; Gaps 1;  
QY 756 CTGGGAGCAGCTCTGGACAGGAGTGGTCCCATCAGGAACCTCGGGCATGGCTGGGAA 815  
Db 10457 CCGGGAGCAGCTCTGGACAGGAGAGTGGTCCCATCCAGGAACCTCGGGCATGGCTGGGAA 10516  
QY 816 GTGGGTACTTGGTCCGGGTCTGTATGTGTGTGACTGGTGTGTGTCAGAGAGAAATGT 875  
Db 10517 GTGGGTACTTGGTCCGGGTCTGTATGTGTGTGACTGGTGTGTGTCAGAGAGAAATGT 10576  
QY 876 GTGCYTAAGTGTACGTGTGAGTCTGTGTATGTGTGTGAATATTGTCTTTGTGGGTGATT 935  
Db 10577 GTGCCCTGAGTGTACGTGTGAGTCTGTGTATGTGTGAATATTGTCTTTGTGGGTGATT 10636  
QY 936 TTCTGCTGTGTAATCGTGTCTCCCTGCAAGTGTGAACAAGTGGACAAAGTGTCTGGGAGTGG 995  
Db 10637 TTCTGCTGTGTAATCGTGTCTCCCTGCAAGTGTGAACAAGTGGACAAAGTGTCTGGGAGTGG 10696  
QY 996 ACAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGCATGTCAAGAGTGCAG 1055  
Db 10697 ACNAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGCATGTCAAGAGTGCAG 10756  
QY 1056 GTGAAGTGAAGGACAGGCCCATGATGCCATCATCATCATCATCATCATCATCATCATCAT 1115  
Db 10757 GTGAAGTGAAGGACAGGCCCATGATGCCATCATCATCATCATCATCATCATCATCATCAT 10816  
QY 1116 TAAAGTCCAGTACACAGATGAAGGTCTGTGGAGGAGGGTGTGTGACTACATTAAGGTGTATGAGCCT 1235  
Db 10817 TAAAGTCCAGTACACAGATGAAGGTCTGTGGAGGAGGGTGTGTGACTACATTAAGGTGTATGAGCCT 1175  
QY 1176 GGAAGGGCAAGGCCATGTCTCTGGAGGAGGGTGTGTGACTACATTAAGGTGTATGAGCCT 1235  
Db 10877 GGAAGGGCAAGGTCTATGTCTGGAGGAGGGTGTGTGACTACATTAAGGTGTATGAGCCT 10936  
QY 1236 AGCTGGAGGTGGATGGCCRGCTCCACTGAAACCCCTGGTTATCCAGAGAGGCTTTTCAGG 1295  
Db 10937 AGCTGGAGGTGGATGGCCGGTCCACTGAGACCCCTGGTTATCCAGAGAGGCTTTTCAGG 10996  
QY 1296 CTTACAGGAGCTTTGGATGGGGAGAGGGGTGACTTCTCCGACAGGCCCTCCACCGGCC 1355  
Db 10997 CTTGGGAGCTTTGGATGGGGAGAGGGGTGACTTCTCCGACAGGCCCTTTCTACACACC 11056  
QY 1356 TACCCTGGCTAAGGGCTCTGGACAGGAAGCAGGGGCAAGAACTCTCTGGAGCAGGCCATAC 1415  
Db 11057 TACCCTGGCTAAGGGCTCTGGACAGGAAGCAGGGGCAAGAACTCTCTGGAGCAGGCCATAC 11116  
QY 1416 CGGCCCTGGCCCTGACTCTG ---CCACTGGCAGCAGTCAACACAGCAGGTTTCATCTACA 1472  
Db 11117 CGGCCCTGGCCCTGACCCCTGCACCCACTGGCAGCAGAGTCAACACAGCAGGTTGGCTACA 11176  
QY 1473 GCAGAGGCAAGGCCATCATCAGTCCCTTTAAGGGAAGGGTCACCGCTCGGTGTG 1532  
Db 11177 GCAGAGGCGAAGGCCATCATCAGTCCCTTTAAGGGAAGGGTCACCGCTCGGTGTG 11236  
QY 1533 CTGAGAGTCTCCTGCTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1592  
Db 11237 CTGAGAGTCTCCTGCTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 11296
```

```
QY 1593 AGGAGCCCATTTGGTAGTGAGGACGATATGGGCTAGAGACACTGGTGCCCTGGCCGTG 1652  
Db 11297 AGGAGCCCATTTGGTAGTGAGGACGATATGGGCTAGAGACACTGGTGCCCTGGCCATG 11356  
QY 1653 ATAGTGGCCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680  
Db 11357 ATAGTGGCCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11384  
RESULT 10  
AX345458 AX345458 5884 bp DNA linear PAT 01-FEB-2002  
LOCUS Sequence 529 from Patent WO0200928.  
DEFINITION AX345458  
ACCESSION AX345458  
VERSION AX345458.1 GI:18493344  
KEYWORDS .  
SOURCE synthetic construct.  
ORGANISM synthetic construct  
1 artificial sequences.  
REFERENCE 1  
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.  
TITLE Diagnosis of diseases associated with the immune system  
JOURNAL Patent: WO 0200928-A 529 03-JAN-2002;  
Epigenomics AG (DE)  
FEATURES  
Location/Qualifiers  
source 1..5884  
/organism="synthetic construct"  
/db_xref="taxon:32630"  
/note="chemically treated genomic DNA (Homo sapiens)"  
BASE COUNT 1259 a 92 c 1514 g 3019 t  
ORIGIN  
Query Match 45.0%; Score 756.4; DB 6; Length 5884;  
Best Local Similarity 70.6%; Pred. No. 2.1e-207;  
Matches 1086; Conservative 5; Mismatches 386; Indels 61; Gaps 4;  
QY 177 TATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGGCACACACCTGTATATCCACGTAC 236  
Db 3551 TAAATTTTTTGTATAAAAAATTAGTATGATGGTGGTGGTATTTGTATTTTACTTTAT 3610  
QY 237 TTAGGAGCTGAGGAGGAGAAATTCCTTCAAGCCTAGAGGTGAAGTTCTAGTGAGCCGA 296  
Db 3611 TTGGGAGGTGAGGTAGGAGAAATCCTTTGAATTTTGGAGGTGGAGGTCTGACTAGTCCA 3670  
QY 297 GATTGCATCATTTGCACAAATGGAGGGAGCCACAGCTGGGCAACAAGAGAAATCTCGG 356  
Db 3671 GATCGTATTATTGTATTTT-----TATTTGGCGGATAGAGTAAATTTCC 3715  
QY 357 TCTCCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 416  
Db 3716 TTTT-----AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3763  
QY 417 ACTTGGGAGGAGGGGTCACCTTGTATGTCGAGACTGACAGTGGGTGGTGGTGGTGGTGGT 476  
Db 3764 ATTGGAGGTAGGAGGTATTATTGATGTTGAGATTGTAGTGTATGATGATGATGATGATG 3823  
QY 477 GCACCTCCGGCCTGGGCAACAGAGTGCACCCCTGCTCTAAAGAA-----AAAAAATAAAG 531  
Db 3824 GTATTTCGGTTCGGGTATAGAGTCAGATTTTGTTTAAGAAAAATAAATAAATAAATAAAG 3883  
QY 532 CAACATATCCTGAACAAAGGATCCTCCATACGTTTCCACACAGATTTCTTAATCAGAAACA 591  
Db 3884 TAAATATATTTAAATAAAGGATTTTATATAATGTTTTTATAGATTTTAAATAGAAATA 3943  
QY 592 TCGAGGCCAGAAAGCAGTGGAGGAGGACACACCTCAGGACGCCCGGGAGGATGTTGTAC 651  
Db 3944 TGGAGGTAGGAAGTAGTGGAGAATGACGATTTTAGTAGTATTTTGGAGGATGTTGTAT 4003  
QY 652 AGGCTGGGCAAGGCCCTTCCGGCTACCAACTGGGAGCTCTGGGAGCCTCGGAGACCCCTGTTGCAA 711  
Db 4004 AGGTTGGGTAAGGGTTTTTAGGTTATTTAATTTGGGAGTTTTCGGGAATAGTTTGTGTGAA 4063  
QY 712 ACAAGAAGCCATAGATCCCGCCAGAGCCAGGA----- 743
```

RESULT 11			
AX348344			
LOCUS	AX348344	5884 bp	DNA
DEFINITION	Sequence 39 from Patent WO0202806.		
ACCESSION	AX348344		
			PAT 06-FEB-2002

```
QY 923 TGTCGGGTGATTTTCGCRFTGCTAATCCTGTCCTCCCTGCAAGTGTGAACAAGTGGACAAG 982
Db 4304 TGTGTGGGTGATTTTGTATGTGTAATCGTGTGTTTGTAAAGTGTGAATAAGTGGATAAG 4363
QY 983 TGTCTGGAGCTGGACAAAGACATCTGTGCACCATCAGCTGTGTCATAGCGTCTGTGCATG 1042
Db 4364 TGTGTGGAGTGGATAGAGATTTGGTATATTAGGTGTGTATAGCGTTTGTGTATG 4423
QY 1043 TCAAGAGTGAAGGTGAAGTGAAGGGACCAAGCCATGATGCCACTCATCATCAGGAGCT 1102
Db 4424 TTAAGAGTGAAGGTGAAGTGAAGGGATTAGTTTATGATGTTATTATTATTAGGAGTT 4483
QY 1103 CTAAAGCCCCAGGTAAAGTGCAGATAGAGGTGCTGAAGGTCACTCTCGAGTGGG 1162
Db 4484 TTAAGGTTTAAAGTAAAGTGTAGTGTAGATAGATAAGGGTGTGAAGGTATTATTGAGTGGG 4543
QY 1163 CAGTGGGGTAGGGAAGGGCAAGGCCATGTTCTGAGGAGGGTGTGCTACTACATTAG 1222
Db 4544 TAGTGGGGTAGGGAAGGGTAGGTTATGTTTGGAGGGGGTGTGATTATATTAG 4603
QY 1223 GGTATGAGCTAGCTGGGAGGTGGATGGCCRGCTCCACTGAAACCCCTGTTATCCAG 1282
Db 4604 GGTATGAGCTTAGTTGGGAGGTGGATGCTCGGTTTATTGAGATTTTGGTTATTAG 4663
QY 1283 AAGCTTTGCAGGCTTCAGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCAGCAGGC 1342
Db 4664 AAGTTTGTGTGGGTTTGGGAGTGTGGAGTGGGAGAGGGGTGATTTTTCGATTAGGT 4723
QY 1343 CCTCCACCGGCTACCTCGGTAAAGGCTTGAGCAGGAGGAGGAGGCAAGACCTCTG 1402
Db 4724 TTTTATTATTATTATTTGGGTAAAGGTTTGGAGTAGGAAGTAGCGGTAAAGGATTTTG 4783
QY 1403 GAGCAGCCCATACCGCCCTGGCTGACTCTGCCACTGGCAGCACAGTCAACACAGCAGG 1462
Db 4784 GAGTAGTTTATTATTGTTTGGTTTGAATTTGTTATTGGTAGTATAGTTAATATAGTAGG 4843
QY 1463 TTCACACAGCAGAGGGCAAGGCCATCATCAGCTCCCTTTTATAAGGAAGGGTCACGC 1522
Db 4844 TTTTATTATAGTAGGGCGAAGGTTATTATTAGTTTATTTTATAAGGAAGGGTTACGC 4903
QY 1523 GCTCGGTGCTCAGAGTGCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1582
Db 4904 GTTCGGTGTCTCAGAGTGTGTTTGTGGTGTGTTTGTGGTGTGTTGTTGGGTGGGTGTAG 4963
QY 1583 GTGTGTCAGAGGAGGCCATTTGGTAGTGAGGAGGTATGGGCTAGAAGCACTGGTGCC 1642
Db 4964 GTGTGTTAGCAGGAGTTTACTTGGTAGTGAGTATGTTATGGGCTAGAAGTATTTGGTT 5023
QY 1643 CTGGGGGTGATAGTGCCCATCTTCCTGCTCCTGGTGG 1680
Db 5024 TTTGGTTATGATAGTGTTATTTTGTGTTTGTGGTGG 5061
```

```
RESULT 12
AX345459/c
LOCUS AX345459 5884 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 530 from Patent WO0200928.
ACCESSION AX345459
VERSION AX345459.1 GI:18493345
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
1.
AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 530 03-JAN-2002;
EpiGenomics AG (DE)
FEATURES
Location/Qualifiers
1..5884
/organism="synthetic construct"
/db_xref="taxon:32630"
```

```
BASE COUNT 1730 a 92 c 1381 g 2681 t
ORIGIN
Query Match 32.0%; Score 537.4; DB 6; Length 5884;
Best Local Similarity 62.4%; Pred. No. 4.5e-144;
Matches 949; Conservative 5; Mismatches 506; Indels 61; Gaps 5;
QY 189 AATAAYAAAAAGCTAGACGTGGTGGCACACACCTGTAAATCCACAGCTACTTATAGGAGGTGA 248
Db 2322 ACAAAAATCAACCAACATAATAATAACACCTATATAATCCCACTACTTAAAAAACTAA 2263
QY 249 GGCAGAGAAATTCCTTGAAGCCCTAGAGGTGAAGTGTGTAGTGAAGCCGAGATTCGATCAT 308
Db 2262 AACAAAAATTCGCTTAAACCTTAAAAATAAAAATACAAATACCAACGAAATCGCATCAT 2203
QY 309 GCACAATGGAGGGAGCCACCGCTGGGCAACAAGAGGAAATCTCGCTCCCAAAAAA 368
Db 2202 ACACCT-----CCACCTAAAGCAGCAAAACAAA-CTCGCTCTC----- 2166
QY 369 AAAAAAATAAAGATTTAGGCTGGGTGGTGTAGTCCAGCTACTTGGGAGGCA 428
Db 2165 -----AAAAAATAATTTAACTCGATAATACCTATATATCCCACTACTTAAAAAACA 2110
QY 429 GGGGTCCACTTGTATGTCGAGACTGCAGTCAGCCATGCTCCTGCCACTGCACCTCCGGCCT 488
Db 2109 AAAAAATCCACTTAAATATTAAATTAACAATAAAGCAATAATCTTACCCTACATCCGACC 2050
QY 489 GGCACAACAGAGTCAGACCTCTCTAAAGAA-----AAAAAATAAAGCAACATATCCTG 543
Db 2049 GAACAACAATAAACCCTTATCTAAAAAATAAATAAATAAACAACATATCCTA 1990
QY 544 AACAAAGGATCCCTCAATAAGTTCCACCACAGATTCTTAATCAGAAACATGAGGCGCCAGAA 603
Db 1989 AATAAAAAATCCCTCCATAATATTTCCACCACAAATTTCTAATCAAAACATAAAAACCAA 1930
QY 604 AGCAGTGGAGGAGACRACCTCAGGAGCGCCGGGAGGATGTTGTGCACAGCTGGGGCAA 663
Db 1929 AACAAATAAAAAATTAAGACCTTCAAAACACCTTAAAAAATACTATCACAACATAAAAACA 1870
QY 664 GGGCTTTCGGCTACCACTGAGGAGCTCTGGGAACAGCCCTGTTGCAACAACAGAGCCAT 723
Db 1869 AAACCTTCAAACTACCACTTAAAACTCTAAACAACACCCCTATTACAAACAATAAATCAT 1810
QY 724 AGCCGCCAGAGCCCGAGGAA-----TGTGGGTGG 754
Db 1809 AACCCGACCAAAACCCCAATAATAAACTAAACTAAATCCATATAAACAATTTAAACTC 1750
QY 755 GCTGGGAGCAGCTCTGGAGAGAGTGGTCCCATCCAGGAAACCTCCGGCATGGCTGGGA 814
Db 1749 ACCGAAACAAACCTCTAAACAAAAAATCCCATCCAAAAAACCCTCGAACATTAACATAA 1690
QY 815 AGTGGGTACTTGGTGGGGTCTGTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 874
Db 1689 AATAAATACTTAATCCGAACTCTATATATATATAAATAAATAAATAAATAAATAAATAAATA 1630
QY 875 TGTGCTTAAGTGTGAGTGTGAGTGTGTATGTGTGAATATATCTCTTTTGTGGGTGAT 934
Db 1629 TATACCTTAATAATCAATAAATCTATATATATAAATAATATCTTTTATATAAATAAT 1570
QY 935 TTTCTGCTGTGTAATCGTCCCTGCAAGTGTGAACAAGTGACAGAGTGTCTGGGAGTG 994
Db 1569 TTTCTACATATATAATCGTATCCCTACAAATATAAACAATAAACAATAAATAAATAAATA 1510
QY 995 GACAAGAGATCTGTGCACCATCAGTGTGTGCATAGCGTCTGTGCATGCTCAAGAGTGCA 1054
Db 1509 AACAAAAATCTATACACCCTCAAAATATATACATAACGCTATACATATACATAAATAAACA 1450
QY 1055 GGTGAAGTGAAGGAGCAGGCCCATGATGCCACTCATCATCAGAGCTCTTAAGGCCCCAG 1114
Db 1449 AATAAATAAATAAACAACCAACCCATTAATACCCTCATCAAAAACCTCTTAACACCCCA 1390
QY 1115 GTAAGTCCAGTGACAGATAAGGGTCTGAAGGTCTACTCTGAGTGGGAGGTGGGGTA 1174
```









```
repeat_region /rpt_family="Alu"
19473..19548 /rpt_family="MER44C"
repeat_region complement(40128..40441)
19689..19985 /rpt_family="Alu"
repeat_region 40790..41056
repeat_region /rpt_family="Alu"
20825..21104 /rpt_family="Alu"
repeat_region 21451..21758
repeat_region /rpt_family="Alu"
complement(21864..22509)
repeat_region /rpt_family="Alu"
22190..22216 /note="(T)27"
repeat_region /rpt_type=tandem
/rpt_unit=T
22670..22693 /note="(T)24"
repeat_region /rpt_type=tandem
/rpt_unit=T
complement(22676..22768)
repeat_region /rpt_family="Alu"
complement(22814..22862)
repeat_region /rpt_family="Alu"
complement(23056..23307)
repeat_region /rpt_family="Alu"
complement(23477..23569)
repeat_region /rpt_family="Alu"
complement(23803..24108)
repeat_region /rpt_family="Alu"
24197..24502
repeat_region /rpt_family="Alu"
24485..24506 /note="(A)22"
repeat_region /rpt_type=tandem
/rpt_unit=A
24904..25177 /rpt_family="Alu"
repeat_region 25370..25510
misc_feature complement(26535..26624)
repeat_region complement(26705..26996)
repeat_region /rpt_family="Alu"
26958..27071 /rpt_family="THE1"
repeat_region complement(28016..28299)
repeat_region /rpt_family="Alu"
complement(28658..28822)
repeat_region /rpt_family="LI"
complement(29152..29437)
misc_feature 31894..32016 /note="GRAIL 2 excellent exon, frame 0"
repeat_region 33793..33935 /rpt_family="Alu"
repeat_region complement(35555..35853)
repeat_region /rpt_family="Alu"
36311..36882 /rpt_family="Alu"
repeat_region 37120..37382 /rpt_family="LI"
repeat_region 38098..38833 /rpt_family="Alu"
repeat_region /rpt_family="MER7"
38352..38639 /rpt_family="Alu"
repeat_region 38823..39113 /rpt_family="Alu"
repeat_region 39195..39433 /rpt_family="Alu"
repeat_region complement(39516..39913)
/rpt_family="LI"

repeat_region 40057..40692 /rpt_family="MER7"
repeat_region complement(40128..40441)
repeat_region 40790..41056
repeat_region /rpt_family="Alu"
42301..42473 /rpt_family="MER58B"
repeat_region 42481..42768

Query Match 17.08; Score 284.8; DB 9; Length 78065;
Best Local Similarity 75.6%; Pred. No. 5.9e-71;
Matches 410; Conservative 3; Mismatches 105; Indels 24; Gaps 4;

QY 1 GAATTCAAGACCAGCCTGGACAACTTGGAAAGAACCSGGTCTCTACAAAAATAACAAAT 60
Db 53829 GAGTTTGAGACCACCTGGGCGGTGCTATGAAACCTTGGGCTCTATAAAAATACAAA-- 53772

QY 61 AGCTGGGATTGGTGGGTGGCTCATGCCCTAATATCCAGCACCTTTGGGAGCCTGAGGTG 120
Db 53771 AAATTAGGCCAGGGGGGTGGCTCATGCCCTGTAATCCAGCACCTTTGGGAGGCCGAGGTG 53712

QY 121 GGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCCTGGCCACATGTTGAACCCCTATC 180
Db 53711 GCGGATCACCTGAGGTTCAGGAGTTCAAGACCAGCCTGACCAACATGGTGAACCCCGTC 53652

QY 181 TCTACTGAAATATYAAAA--AGCTAGACGTGGTGGCACACACCTGTAATCCAGCTACTT 238
Db 53651 TCTACTGAAATATYAAAA--AGCTAGACGTGGTGGCACACACCTGTAATCCAGCTACTT 53592

QY 239 AGGAGCTGAGGCGAGGAGATTCCTTGAAGCCTTAGAGTGAAGTGTAGTGAAGCCGAGA 298
Db 53591 GGGAGCTGAGGCGAGGAGATTCCTTGAAGCCTTAGAGTGAAGTGTAGTGAAGCCGAGA 53532

QY 299 TTGCATCATTCGACAAATGGAGGGAGCCACAGCCTGGGCAACAGAGAAATCTCCGTC 358
Db 53531 TTGCATCATTCGACAAATGGAGGGAGCCACAGCCTGGGCAACAGAGAAATCTCCGTC 53486

QY 359 TCACAAAAAAMAAAAAAMAAAAAAGRATTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 418
Db 53485 TCACAAAAAAMAAAAAAMAAAAAATATCTGGGCGTGGTGGTGGTGGTGGTGGTGGTGGT 53426

QY 419 TTGGGAGGC-----AGGGGTCCACTTCATGTCGAGACTGCAGTCAGCCATGATCCTGC 472
Db 53425 TTGGGAGGC-----AGGGGTCCACTTCATGTCGAGACTGCAGTCAGCCATGATCCTGC 53366

QY 473 CACTGACTCCGGCCTGGGCAACAGAGTGAAGCCTGTCTAAAGAAAAAATAAAGC 532
Db 53365 CACTGACTCCGGCCTGGGCAACAGAGTGAAGCCTGTCTAAAGAAAAAATAAAGC 53306

QY 533 AA 534
Db 53305 AA 53304

RESULT 15
AC087389 LOCUS
DEFINITION Homo sapiens chromosome 17 clone RP11-304114 map 17, WORKING DRAFT
SEQUENCE, 7 unordered pieces.
ACCESSION AC087389
VERSION AC087389.2 GI:22381544
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 191280)
AUTHORS Birren,B., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone RP11-304114
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 191280)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
```





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:43:59 ; Search time 682.639 Seconds  
(without alignments)  
5542.256 Million cell updates/sec

Title: US-09-942-310-2  
Perfect score: 1680  
Sequence: 1 gaattcaagaccgctgga.....catcttcgtctcgtgtgg 1680

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1677.2	99.8	AAD34214	Human CYP2D6 gene
2	1677.2	99.8	AAD34213	Human cytochrome P
3	1665.4	99.1	AAH26169	Human cytochrome P
4	1665.4	99.1	AAH26179	Human cytochrome P
5	1059.4	63.1	ABQ72364	Human CYP2D6 gene,
6	1056.6	62.9	ABQ72215	Human CYP2D6 gene,
7	756.4	45.0	ABK39958	Human chemically p
8	756.4	45.0	ABL32556	Human immune syste
9	537.4	32.0	ABL32557	Human immune syste

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

c	10	537.4	32.0	5884	24	ABL32557	Human immune syste
	11	258.2	15.4	122888	24	ABK83569	Human cDNA differe
	12	254.4	15.1	35959	22	AAK78275	Human immune/haema
	13	253.6	15.1	129722	24	ABQ88117	Human osteoblast d
	14	252.6	15.0	13273	22	AAK89658	Human digestive sy
	15	252	15.0	32169	22	ABA14358	Human nervous syst
	16	251.8	15.0	86080	24	ABQ88164	Human osteoblast d
	17	251.8	15.0	86080	24	ABK83561	Human cDNA differe
	18	250	14.9	10684	17	AAK33758	Control region iso
	19	249.6	14.9	220895	24	ABK84798	Human cDNA differe
	20	249.2	14.8	13409	22	ABA08135	Human ovarian and
	21	249.2	14.8	13409	22	AAK06913	Human reproductive
	22	247.6	14.7	3533	22	AAK94890	Human full-length
	23	247	14.7	5446	22	AAI199040	Human excretory re
	24	247	14.7	5446	22	AAI163390	Human kidney relat
c	25	244.4	14.5	40668	24	ABQ88150	Human osteoblast d
c	26	244.2	14.5	26427	22	ABA20762	Human nervous syst
c	27	244.2	14.5	26427	22	ABA20763	Human nervous syst
	28	243.8	14.5	5881	22	ABL07230	Human reproductive
	29	243.8	14.5	5881	23	ABL98778	Human testicular a
c	30	243	14.5	9192	22	AAS33461	DNA encoding human
c	31	243	14.5	25715	22	AAS33462	Human ovarian and
	32	242.2	14.4	13862	22	ABA08208	Human reproductive
	33	242.2	14.4	13862	22	AAI02789	Human reproductive
	34	242.2	14.4	13862	22	AAI07516	Nucleotide sequenc
	35	241.6	14.4	110000	22	AAF84800	Human osteoblast d
c	36	241.4	14.4	92638	24	ABQ88096	Human nervous syst
c	37	240.6	14.3	10663	22	ABA18961	Human immune/haema
c	38	240.6	14.3	37314	22	AAK71358	Human osteoblast d
	39	240.6	14.3	172570	24	ABQ88207	Human heparin cofa
c	40	240.2	14.3	15848	20	AAZ32190	Gene #2362 used to
c	41	240.2	14.3	15849	24	ABN95864	Genomic sequence #
c	42	240.2	14.3	21470	23	ABK42270	Human immune/haema
	43	239.8	14.3	19659	22	AAK64731	Human immune/haema
	44	239.8	14.3	19859	22	AAK75600	Human digestive sy
	45	239.8	14.3	19659	22	AAK91504	Human digestive sy

ALIGNMENTS

RESULT 1  
AAD34214  
ID AAD34214 standard; DNA: 1680 BP.  
AC AAD34214;  
XX  
XX  
DT 16-JUL-2002 (first entry)  
XX  
DE Human CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX  
KW Human: cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; polymorphism; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT misc\_feature 36  
FT Location/Qualifiers  
FT misc\_feature 194  
FT /tag= a  
FT /note= "Polymorphic site"  
FT /tag= b  
FT /note= "Polymorphic site"  
FT misc\_feature 385  
FT /tag= c  
FT /note= "Polymorphic site"  
FT misc\_feature 620  
FT /tag= d  
FT /note= "Polymorphic site"  
FT misc\_feature 880  
FT /tag= e  
FT /note= "Polymorphic site"

FT misc\_feature 942 /\*tag= f  
FT /\*note= "Polymorphic site"  
FT 1255  
FT /\*tag= g  
FT /\*note= "Polymorphic site"  
XX  
PN WO200218638-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 27-AUG-2001; 2001WO-IB01544.  
XX  
PR 30-AUG-2000; 2000GB-0021286.  
XX  
XX (GEMI-) GEMINI GENOMICS PLC.  
XX  
XX Risinger C, Andersson MK, Lewander T, Oliasson E;  
XX WPI; 2002-329785/36.  
XX  
XX New sequence determination oligonucleotides, useful for detecting  
PT polymorphic sites in a 5' flanking region of a Cyp2d6 gene, as  
PT hybridization probes, as components of diagnostic assays, or in  
PT ligase-based sequence determination  
XX  
XX Claim 1; Fig 2; 63pp; English.  
XX  
XX The invention relates to sequence determination oligonucleotides for  
CC detecting polymorphic sites in a 5' flanking region of cytochrome P450  
CC 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many  
CC different xenobiotics. Human CYP2D6 gene is located on chromosome 22.  
CC The oligonucleotides may be used as in situ hybridisation probes, in  
CC ligase-based sequence determination, as components of diagnostic assays,  
CC as probes in sequence determination methods based on mismatches, as  
CC hybridisation-based diagnostic assays, and as components of diagnostic  
CC microarray. CYP2D6 is useful to predict variations in an individual's  
CC ability to metabolise certain drugs. The present sequence is human  
CC CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX  
XX Sequence 1680 BP; 413 A; 379 C; 539 G; 342 T; 7 other;  
XX  
Query Match 99.8%; Score 1677.2; DB 24; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAATTCAGACCCCTGGACACTTGGGAAGAACCSGGTCTCTACAAAAAATACAAAAAT 60  
Db |  
1 GAATTCAGACCCCTGGACACTTGGGAAGAACCSGGTCTCTACAAAAAATACAAAAAT 60  
QY 61 AGCTGGGATTGGGTGGGTCATGCCCTATTAATCCACGACTTTGGGAGCCTGAGGTG 120  
Db |  
61 AGCTGGGATTGGGTGGGTCATGCCCTATTAATCCACGACTTTGGGAGCCTGAGGTG 120  
QY 121 GGTGGATCACCTGAAGTCAGGAGTCAAGACTAGCCTGGCCAAACATGGTGAACCCCTATC 180  
Db |  
121 GGTGGATCACCTGAAGTCAGGAGTCAAGACTAGCCTGGCCAAACATGGTGAACCCCTATC 180  
QY 181 TCTACTGAANAATAYAAAAAGCTAGAGTGGGCACACACCTGTAAATCCAGCTACTTAG 240  
Db |  
181 TCTACTGAANAATAYAAAAAGCTAGAGTGGGCACACACCTGTAAATCCAGCTACTTAG 240  
QY 241 GAGGCTGAGGCAGGAGAAATGCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCCGAGATT 300  
Db |  
241 GAGGCTGAGGCAGGAGAAATGCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCCGAGATT 300  
QY 301 GCATCATTCACAATAGAGGGAGGCCACACGCTGGGCAACAGAGAAATCTCCGTCAC 360  
Db |  
301 GCATCATTCACAATAGAGGGAGGCCACACGCTGGGCAACAGAGAAATCTCCGTCAC 360  
QY 361 CAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 420  
Db |  
361 CAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 420

QY 421 GGGAGGAGGGGGTCCACTTGATGTCGAGACTGCAGTGAAGCCATGATCCTGCGCACTGCAC 480  
Db |  
421 GGGAGGAGGGGGTCCACTTGATGTCGAGACTGCAGTGAAGCCATGATCCTGCGCACTGCAC 480  
QY 481 TCCGGCTGGGCAACAGAGTGAACCCCTGTCTAAAGAAAAAATAAAGCAACATATC 540  
Db |  
481 TCCGGCTGGGCAACAGAGTGAACCCCTGTCTAAAGAAAAAATAAAGCAACATATC 540  
QY 541 CTGAACAAAGATCTCCATTAACGTTCCACCGAGTTTCTAATCAGAAACATGAGGCCA 600  
Db |  
541 CTGAACAAAGATCTCCATTAACGTTCCACCGAGTTTCTAATCAGAAACATGAGGCCA 600  
QY 601 GAAAGCAGTGGAGGAGACRACCTCAGGAGCCCGGAGGATGTTCTCAGAGCTGGG 660  
Db |  
601 GAAAGCAGTGGAGGAGACRACCTCAGGAGCCCGGAGGATGTTCTCAGAGCTGGG 660  
QY 661 CAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTGCAACAAGAGC 720  
Db |  
661 CAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTGCAACAAGAGC 720  
QY 721 CATAGCCCGCCAGAGCCAGGAATGTGGCTGGGCTGGGAGAGCCTCTGGACAGAGT 780  
Db |  
721 CATAGCCCGCCAGAGCCAGGAATGTGGCTGGGCTGGGAGAGCCTCTGGACAGAGT 780  
QY 781 GGTCCCATCCAGGAAACCTCCGGCATGGCTGGGAAGTGGGGTACTTGGTCCCGGGTCTGT 840  
Db |  
781 GGTCCCATCCAGGAAACCTCCGGCATGGCTGGGAAGTGGGGTACTTGGTCCCGGGTCTGT 840  
QY 841 ATGTGTGTGACTGTGTGTGTGTGAGAGAGAAATGTGTGCTAAGTGTGAGTGTGAGTCT 900  
Db |  
841 ATGTGTGTGACTGTGTGTGTGTGAGAGAGAAATGTGTGCTAAGTGTGAGTGTGAGTCT 900  
QY 901 GTCTATGTGTGAATATGTCTTGTGTGGTGTATTTCTGCTGTGTATCTGCTGTGCTG 960  
Db |  
901 GTCTATGTGTGAATATGTCTTGTGTGGTGTATTTCTGCTGTGTATCTGCTGTGCTG 960  
QY 961 CAAAGTGTGAACAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGT 1020  
Db |  
961 CAAAGTGTGAACAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGT 1020  
QY 1021 GTGTGATAGCGTCTGTGATGTCAGAGTGAAGTGAAGTGAAGGACAGGCCCCATG 1080  
Db |  
1021 GTGTGATAGCGTCTGTGATGTCAGAGTGAAGTGAAGTGAAGGACAGGCCCCATG 1080  
QY 1081 ATGCCACTCATCATCAGAGAGCTCTAAGGCCCCAGGTGAAGTGAAGTGAAGGACAGG 1140  
Db |  
1081 ATGCCACTCATCATCAGAGAGCTCTAAGGCCCCAGGTGAAGTGAAGTGAAGGACAGG 1140  
QY 1141 CTGAAGGTCACCTCTGGAGTGGGAGTGGGAGTGGGAGGAAAGGCAAGGCCATGTTCTGA 1200  
Db |  
1141 CTGAAGGTCACCTCTGGAGTGGGAGTGGGAGTGGGAGGAAAGGCAAGGCCATGTTCTGA 1200  
QY 1201 GGAGGGTGTGCTACATACATTAAGGTGTATAGACCTAGCTGGGAGTGGATGCCRGGTCC 1260  
Db |  
1201 GGAGGGTGTGCTACATACATTAAGGTGTATAGACCTAGCTGGGAGTGGATGCCRGGTCC 1260  
QY 1261 ACTGAAACCCCTGTTATCCCAAGGCTTTGACAGGCTTCAGAGCTTCAGTGGGAGAG 1320  
Db |  
1261 ACTGAAACCCCTGTTATCCCAAGGCTTTGACAGGCTTCAGAGCTTCAGTGGGAGAG 1320  
QY 1321 GGGGTGACTTCTCCAGCAGGCCCTCCACCGGCCCTACCTGGGTGAAGGGCTGAGCAG 1380  
Db |  
1321 GGGGTGACTTCTCCAGCAGGCCCTCCACCGGCCCTACCTGGGTGAAGGGCTGAGCAG 1380  
QY 1381 GAAAGCAGGGCAAGAACCTCTGGAGCAGCCCATACCCGCCCTGAGTGTGAGGAGT 1440  
Db |  
1381 GAAAGCAGGGCAAGAACCTCTGGAGCAGCCCATACCCGCCCTGAGTGTGAGGAGT 1440  
QY 1441 GCAGCAGTCAACACAGCAGGTTTCACTCAGAGAGGGCAAGGCCATCATCAGTCC 1500  
Db |  
1441 GCAGCAGTCAACACAGCAGGTTTCACTCAGAGAGGGCAAGGCCATCATCAGTCC 1500

QY	1501	CTTTATAAGGAAGGTCACGGCTCGGTGCTGCTGAGAGTGTCTCTGCCTGGTCTCTGTG	1560
Db	1501	CTTTATAAGGAAGGTCACGGCTCGGTGCTGCTGAGAGTGTCTCTGCCTGGTCTCTGTG	1560
QY	1561	CTCGTGGGTGGGGTGGCCAGGTGTCTCAGAGGAGCCCATTTGGTGTAGTGAGGCAGGTA	1620
Db	1561	CTCGTGGGTGGGGTGGCCAGGTGTCTCAGAGGAGCCCATTTGGTGTAGTGAGGCAGGTA	1620
QY	1621	TGGGGCTAGAACGACACTGGTGCCCTGGCCGTGATAGTGGCCATCTTCTCTCTGGTGG	1680
Db	1621	TGGGGCTAGAACGACACTGGTGCCCTGGCCGTGATAGTGGCCATCTTCTCTCTGGTGG	1680
RESULT 2			
AAD34213			
ID	AAD34213 standard; DNA; 9432 BP.		
XX			
AC	AAD34213;		
XX			
DT	16-JUL-2002 (first entry)		
XX	Human cytochrome P450 2D6 (CYP2D6) gene.		
DE			
XX	Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic; ligase-based sequenced determination; drug metabolism; chromosome 22; gene; ds.		
KW			
KW			
KW			
XX			
OS	Homo sapiens.		
XX			
PN	WO200218638-A2.		
XX			
PD	07-MAR-2002.		
XX			
PF	27-AUG-2001; 2001WO-IB01544.		
XX			
PR	30-AUG-2000; 2000GB-0021286.		
XX			
XX	(GEMI-) GEMINI GENOMICS PLC.		
PA			
XX			
PI	Rislinger C, Andersson MK, Lewander T, Oliasson E;		
XX			
DR	WPI; 2002-329785/36.		
XX			
PT	New sequence determination oligonucleotides, useful for detecting polymorphic sites in a 5' flanking region of a CYP2D6 gene, as hybridization probes, as components of diagnostic assays, or in ligase-based sequence determination -		
PT			
PT			
XX	Example 3; Fig 1; 63pp; English.		
PS			
XX			
CC	The invention relates to sequence determination oligonucleotides for detecting polymorphic sites in a 5' flanking region of cytochrome P450 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many different xenobiotics. Human CYP2D6 gene is located on chromosome 22. The oligonucleotides may be used as in situ hybridisation probes, in ligase-based sequenced determination, as components of diagnostic assays, as probes in sequence determination methods based on mismatches, as hybridisation-based diagnostic assays, and as components of diagnostic microarray. CYP2D6 is useful to predict variations in an individual's ability to metabolise certain drugs. The present sequence is human CYP2D6 gene.		
CC			
XX			
SQ	Sequence 9432 BP; 1964 A; 2647 C; 2976 G; 1845 T; 0 other;		
Query Match 99.8%; Score 1677.2; DB 24; Length 9432;			
Best Local Similarity 99.6%; Pred. No. 0;			
Matches 1673; Conservative 7; Mismatches 0; Indels 0; Gaps 0;			
QY	1	GAATTCAAGACCAAGCCCTGGACAACTTGGGAAGAACCSGGTCTCTACAAAAATACAAAAAT	60
Db	1	GAATTCAAGACCAAGCCCTGGACAACTTGGGAAGAACCSGGTCTCTACAAAAATACAAAAAT	60
QY	61	AGCTGGGATTGGGTGGGTGGTGCATGCCTATATAATCCAGCACTTTGGGAGCCTGAGGTG	120

Db	61	AGCTGGGATTGGGTGGGTGGTGCATGCCTATATAATCCAGCACTTTGGGAGCCTGAGGTG	120
QY	121	GGTGGATCACCTCAAGTCAAGAGTTCAAAGACTAGCCTGGCCAAACATGGTGAACCCCTATC	180
Db	121	GGTGGATCACCTCAAGTCAAGAGTTCAAAGACTAGCCTGGCCAAACATGGTGAACCCCTATC	180
QY	181	TCTACTGAAAAATAYAAAAAGCTAGACGTGGTGGCACACACCTCTAATCCACAGCTACTTAG	240
Db	181	TCTACTGAAAAATACAAAAAGCTAGACGTGGTGGCACACACCTCTAATCCACAGCTACTTAG	240
QY	241	GAGGCTGAGGCAGGAGAAATTCCTTGAAGCCTAGAGCTGAAGTTCCTAGTGAGCCGAGATT	300
Db	241	GAGGCTGAGGCAGGAGAAATTCCTTGAAGCCTAGAGTGAAGTTCCTAGTGAGCCGAGATT	300
QY	301	GCATCATTCACAAATGGAGGGAGCCACAGCCTGGCCAAACAGAGAAATCTCCGTCTC	360
Db	301	GCATCATTCGACAATGGAGGGAGCCACAGCCTGGCCAAACAGAGAAATCTCCGTCTC	360
QY	361	CAAAAAAAGGATTCCTCCATAACGTTCCACACAGATTTCTAATCAGAAACATGATATC	420
Db	361	CAAAAAAAGGATTCCTCCATAACGTTCCACACAGATTTCTAATCAGAAACATGATATC	420
QY	421	GGAGGAGGGGGTCCACTTGATGTCGAGACTGCAGTGAGCCATGATCCTGCCACTGCAC	480
Db	421	GGAGGAGGGGGTCCACTTGATGTCGAGACTGCAGTGAGCCATGATCCTGCCACTGCAC	480
QY	481	TCCGGCTTGGGCAACAGAGTGAGACCCCTGTCTAAAGAAAAAATAAAGCAACATATC	540
Db	481	TCCGGCTTGGGCAACAGAGTGAGACCCCTGTCTAAAGAAAAAATAAAGCAACATATC	540
QY	541	CTGAACAAAGGATTCCTCCATAACGTTCCACACAGATTTCTAATCAGAAACATGAGGCCA	600
Db	541	CTGAACAAAGGATTCCTCCATAACGTTCCACACAGATTTCTAATCAGAAACATGAGGCCA	600
QY	601	GAAGCAGTGGAGGAGGACRACCCCTCAGGCAGCCCCGGGAGGATGTTGTACAGCGCTGGG	660
Db	601	GAAGCAGTGGAGGAGGACRACCCCTCAGGCAGCCCCGGGAGGATGTTGTACAGCGCTGGG	660
QY	661	CAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTTGCAACAAAGAAC	720
Db	661	CAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGAACAGCCCTGTTGCAACAAAGAAC	720
QY	721	CATAGCCCGGCCAGAGCCAGGAATGTGGCTGGGCTGGGAGCAGCTCTGGACAGAGT	780
Db	721	CATAGCCCGGCCAGAGCCAGGAATGTGGCTGGGCTGGGAGCAGCTCTGGACAGAGT	780
QY	781	GGTCCCATCCAGGAACCTCCGGCATGGCTGGGAAGTGGGGTACTTGGTCCCGGCTCTGT	840
Db	781	GGTCCCATCCAGGAACCTCCGGCATGGCTGGGAAGTGGGGTACTTGGTCCCGGCTCTGT	840
QY	841	ATGTGTGTGTGACTGTGTGTGTGAGAGAGATGTCTGCYCTAAGTGTCAAGTGTGAGTCT	900
Db	841	ATGTGTGTGTGACTGTGTGTGTGAGAGAGATGTCTGCYCTAAGTGTCAAGTGTGAGTCT	900
QY	901	GTGTATGTGTGAATATTGTTTGTGTGGGTGATTTTCTGCRGTGTGTAAATCGTCCCTG	960
Db	901	GTGTATGTGTGAATATTGTTTGTGTGGGTGATTTTCTGCRGTGTGTAAATCGTCCCTG	960
QY	961	CAAGTGTGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGT	1020
Db	961	CAAGTGTGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGT	1020
QY	1021	GTGTGCATAGCGTCTGTGCATGTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	1080
Db	1021	GTGTGCATAGCGTCTGTGCATGTCAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT	1080
QY	1081	ATGCCACTCATCATCAGAGCTCTTAAGGGCCCCCAGAGTAAGTGCACAGTGAAGGGTG	1140
Db	1081	ATGCCACTCATCATCAGAGCTCTTAAGGGCCCCCAGAGTAAGTGCACAGTGAAGGGTG	1140
QY	1141	CTGAAGGTCACTCTTGAGTGGGAGGTGGGGTGAAGGAAAGGCAAGGCCATGTTCTTGGAA	1200
Db	1141	CTGAAGGTCACTCTTGAGTGGGAGGTGGGGTGAAGGAAAGGCAAGGCCATGTTCTTGGAA	1200

```
Db 1141 CTGAAGGTCACCTGGAGTCGGCAGGTGGGGTAGGGAAGGCAAGGCCATGTTCTGGA 1200
QY 1201 GGAGGGTTGTGACTACATATTAGGTGTATGAGCTAGCTGGAGGTGCATGCCRGGTCC 1260
Db 1201 GGAGGGTTGTGACTACATATTAGGTGTATGAGCTAGCTGGAGGTGCATGCCRGGTCC 1260
QY 1261 ACTGAACCCCTGGTTATCCCAAGAGGCTTTTCAGAGCTTCAGAGCTTGGAGTGGGAGAG 1320
Db 1261 ACTGAACCCCTGGTTATCCCAAGAGGCTTTTCAGAGCTTCAGAGCTTGGAGTGGGAGAG 1320
QY 1321 GGGGTGACTTCTCGACACAGGCCCTCCACGGGCTTACCTGGTAAAGGGCCTGGAGCAG 1380
Db 1321 GGGGTGACTTCTCGACACAGGCCCTCCACGGGCTTACCTGGTAAAGGGCCTGGAGCAG 1380
QY 1381 GAAGCAGGGGCAAGAACCTCTGGAGCAGCCCATACCGCCCTGCGCTGACTCTGCCACTG 1440
Db 1381 GAAGCAGGGGCAAGAACCTCTGGAGCAGCCCATACCGCCCTGCGCTGACTCTGCCACTG 1440
QY 1441 GCAGCAGCTCAACACAGCAGGTTCACTACAGCAGAGGCAAGGCCATCATCAGCTCC 1500
Db 1441 GCAGCAGCTCAACACAGCAGGTTCACTACAGCAGAGGCAAGGCCATCATCAGCTCC 1500
QY 1501 CTTTATAAGGGAAGGTCACGCGCTCGGTGCTGAGAGTGTCTCGCTGGTCTCTGTG 1560
Db 1501 CTTTATAAGGGAAGGTCACGCGCTCGGTGCTGAGAGTGTCTCGCTGGTCTCTGTG 1560
QY 1561 CCTGGTGGGTGGGGTGCCAGTGTGTCCAGAGGAGCCCATTTGTTAGTGAGGAGGTA 1620
Db 1561 CCTGGTGGGTGGGGTGCCAGTGTGTCCAGAGGAGCCCATTTGTTAGTGAGGAGGTA 1620
QY 1621 TGGGGCTAGAGAGCACTGGTGCCTGCGCTGATAGTGGGCCATCTTCTGCTCTGTGGT 1680
Db 1621 TGGGGCTAGAGAGCACTGGTGCCTGCGCTGATAGTGGGCCATCTTCTGCTCTGTGGT 1680

RESULT 3
AAH26169
ID AAH26169 standard; DNA; 1669 BP.
XX
AC AAH26169;
CC
DT 17-SEP-2001 (first entry)
XX
DE Human cytochrome P450 CYP2D6 gene promoter region.
XX
KW Cytochrome P450; CYP2D6; promoter; drug metabolism; human;
KW diagnosis; therapy; ds.
XX
OS Homo sapiens.
XX
FH Key
FT primer_bind
FT complement (14..36)
FT /tag= a
FT /note= "amplification primer upf14"
FT primer_bind
FT 337..358
FT /tag= b
FT /note= "sequence primer R1"
FT primer_bind
FT 493..514
FT /tag= c
FT /note= "sequencing primer R2"
FT primer_bind
FT complement (565..577)
FT /tag= d
FT /note= "sequencing primer F2"
FT primer_bind
FT 602..620
FT /tag= e
FT /note= "sequencing primer R3"
FT primer_bind
FT complement (968..988)
FT /tag= f
FT /note= "sequencing primer F3"
FT primer_bind
FT 1124..1143
FT /tag= g
FT /note= "sequencing primer R4"
FT primer_bind
FT 1605..1623
```

```
FT /tag= h
FT /note= "sequencing primer R5"
FT primer_bind
FT 1650..1669
FT /tag= i
FT /note= "amplification primer upr1669"
FT 5'UTR
FT 1532..1619
FT /tag= j
FT CDS
FT 1620..1669
FT /tag= k
FT /partial
FT /note= "5' region of CYP2D6 coding region"
PN WO200155432-A2.
XX
XX 02-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-EP00954.
XX
XX 31-JAN-2000; 2000EP-0101889.
XX
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX
XX Raimundo S, Zanger U;
PI
XX
XX WPI; 2001-457734/49.
XX
XX A polynucleotide capable of hybridizing to CYP2D6 promoter useful for
the optimization of drug therapies using substrates of cytochrome P-450
_
Claim 1; Fig 1; 41pp; English.
CC
CC The present sequence is that of the promoter region of the human
CC cytochrome P450 CYP2D6 gene. The promoter region was amplified
CC by PCR from leucocyte DNA of over 50 individuals, and sequenced.
CC 8 Previously unknown single nucleotide polymorphisms (SNP) were
CC identified. These were at: base 36 (base -1584 according to the
CC Human Cytochrome P450 Allele Nomenclature), where the SNP was C to
CC G, occurring at an estimated frequency of approximately 20% in the
CC whole population, and resulting in increased enzyme activity;
CC position 194 (-1426), C to T, approximately 20% frequency, neutral
CC function; position 385 (-1235), A to G, approximately 50% frequency,
CC neutral function; position 620 (-1000), G to A, approximately 20%
CC frequency, neutral function; position 880 (-740), C to T,
CC approximately 30% frequency, unknown function; position 940 (-580),
CC G to A, approximately 30% frequency, unknown function; 1255 (-365),
CC G to A, rare, unknown function; and 1298 (-322), T to C, rare,
CC unknown function. The C to G mutation at -1584 bp is strongly
CC associated with lower metabolic ratios, and a molecular variant
CC polynucleotide having G at this position is claimed (see AAH26179).
CC The invention provides a method of diagnosing a disorder related to
CC reduced or enhanced capacity for clearance of CYP2D6 substrates
CC (antiarrhythmic, beta-adrenergic receptor-antagonist, tricyclic
CC antidepressant, selective serotonin reuptake inhibitor, neuroleptic,
CC opiates, cytostatic or amphetamine), or susceptibility to such a
CC disorder, by determining the presence of a mutation in the CYP2D6
CC promoter. The strong association of the common C to G mutation at
CC -1584 bp with increased enzyme activity significantly improves the
CC correlation between genotype and phenotype in the CYP2D6 polymorphism.
CC Testing for the mutation will allow the identification of intermediate
CC metabolizers and therefore allow quantitative predictions to be made
CC on in vivo drug metabolism capacity, thus providing a very potent
CC tool for improving the therapy of diseases with drugs that are
CC targets of the CYP2D6 gene product.
XX
SQ Sequence 1669 BP; 413 A; 376 C; 534 G; 338 T; 8 other;
Query Match 99.1%; Score 1665.4; DB 22; Length 1669;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1666; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAATTTCAGACCCCTGGACAACTTGGCAAGAACCSGCTCTCTACAAAAATACAAATT 60
|||||
```



Db	1	CAATTCAAGACACGACCTGGGACAACTTTGGGAAGAACCSGGTCTCTACAAAAATACAAAATT	60
Qy	61	AGCTGGGATTTGGGTGGGGTGCATCCCTATAATCCAGCACACTTTGGAGACCTCAGGTG	120
Db	61	AGCTGGGATTTGGGTGGGGTGCATCCCTATAATCCAGCACACTTTGGAGACCTCAGGTG	120
Qy	121	GGTGGATCACCCTGAAGTCAGGAGTTCAAGACTAGCCTGGCGAACATGTTGAAACCCCTATC	180
Db	121	GGTGGATCACCCTGAAGTCAGGAGTTCAAGACTAGCCTGGCGAACATGTTGAAACCCCTATC	180
Qy	181	TCCTACTGAAATATYAAAAAGCTAGAGCTGGTGGSCACACACTGTAAATCCAGCTACTCTAG	240
Db	181	TCCTACTGAAATATYAAAAAGCTAGAGCTGGTGGSCACACACTGTAAATCCAGCTACTCTAG	240
Qy	241	GAGGCTGAGCGAGGAGAAATGCTTGAAGCCTAGAGGTTGAAGGTTGTAGTGAGCCGAGATT	300
Db	241	GAGGCTGAGCGAGGAGAAATGCTTGAAGCCTAGAGGTTGAAGGTTGTAGTGAGCCGAGATT	300
Qy	301	GCATCATTTGCACAATGGAGGGAGGCCACAGCCTGGGCAACAAGAGGAAATCTCCGTTCTC	360
Db	301	GCATCATTTGCACAATGGAGGGAGGCCACAGCCTGGGCAACAAGAGGAAATCTCCGTTCTC	360
Qy	361	CAAAAAAAAAAAAAAAAAAGATTAAGGCTGGTGGTGCCTGTAGTCCCGACTACTT	420
Db	361	CAAAAAAAAAAAAAAAAAAGATTAAGGCTGGTGGTGCCTGTAGTCCCGACTACTT	420
Qy	421	GGGAGCGAGGGGTCCTCACTTGATGTCGAGACTGCAGTCAGGCATGATCCTGCCACTGCAC	480
Db	421	GGGAGCGAGGGGTCCTCACTTGATGTCGAGACTGCAGTCAGGCATGATCCTGCCACTGCAC	480
Qy	481	TCGGGCTGGGCAACAGAGTGAGACCTGCTTAAAGAAAAAATAAAGCAACATATC	540
Db	481	TCGGGCTGGGCAACAGAGTGAGACCTGCTTAAAGAAAAAATAAAGCAACATATC	540
Qy	541	CTGAACAAGAGATCCTCCATAAGTTCCACCAGATTTCTTAATCAGAACATGAGAGCCA	600
Db	541	CTGAACAAGAGATCCTCCATAAGTTCCACCAGATTTCTTAATCAGAACATGAGAGCCA	600
Qy	601	GAAGCAGTGGAGGAGGACRACCTCAGGACGCCGGGAGGATGTTGTACAGGCTGGGG	660
Db	601	GAAGCAGTGGAGGAGGACRACCTCAGGACGCCGGGAGGATGTTGTACAGGCTGGGG	660
Qy	661	CAAGGGCTTCCGGCTTACCACTGGGAGCTTGGGAACAGCCCTGTTGCAAAACAAGAAC	720
Db	661	CAAGGGCTTCCGGCTTACCACTGGGAGCTTGGGAACAGCCCTGTTGCAAAACAAGAAC	720
Qy	721	CATAGCCCGGCAGACCCAGGAATGTGGGCTGGGCTGGGAGCAGCCTCTGGACAGAGT	780
Db	721	CATAGCCCGGCAGACCCAGGAATGTGGGCTGGGCTGGGAGCAGCCTCTGGACAGAGT	780
Qy	781	GGTCCCATCCAGGAAACCTCCGGCATGGCTGGGAAGTGGGGTACTTGTGTCCGGGCTCTG	840
Db	781	GGTCCCATCCAGGAAACCTCCGGCATGGCTGGGAAGTGGGGTACTTGTGTCCGGGCTCTG	840
Qy	841	ATGTGTGTGTACTGGTGTGTGTGAGAGAGAAATGTGTGCYCTAAGTGTCTAGTGTGAGTCT	900
Db	841	ATGTGTGTGTACTGGTGTGTGTGAGAGAGAAATGTGTGCYCTAAGTGTCTAGTGTGAGTCT	900
Qy	901	GTGTATGTGTGAATATCTTTTGTGTGGGTGATTTTCTGCTGTGTAACTGCTGTCCCTG	960
Db	901	GTGTATGTGTGAATATCTTTTGTGTGGGTGATTTTCTGCTGTGTAACTGCTGTCCCTG	960
Qy	961	CAAGTGTGAACAAGTGCACAGTGTCTGGGAGTGCACAAGAGATCTGTGCACCATCAGGT	1020
Db	961	CAAGTGTGAACAAGTGCACAGTGTCTGGGAGTGCACAAGAGATCTGTGCACCATCAGGT	1020
Qy	1021	GTGTGCATAGCGTCTGTGCATGTCAAGAGTGCAGAGTGAAGTGAAGGACACAGGCCCATG	1080
Db	1021	GTGTGCATAGCGTCTGTGCATGTCAAGAGTGAAGTGAAGGACACAGGCCCATG	1080
Qy	1081	ATGCCACTCATCATCAGGAGCTCTTAAGGCCCCAGGTAAAGTGCCTAGTGACAGATAAGGGT	1140
Db	1081	ATGCCACTCATCATCAGGAGCTCTTAAGGCCCCAGGTAAAGTGCCTAGTGACAGATAAGGGT	1140

Qy	1141	CTGAAGGTCACTCTGGAGTGGGCAGGTGGGGGTAGGGAAAGGCCAAGTGTCTTGGA	1200
Db	1141	CTGAAGGTCACTCTGGAGTGGGCAGGTGGGGGTAGGGAAAGGCCAATGTTCTGGA	1200
Qy	1201	GGAGGGTTCGTGACTACATATTAGGCTGTATGAGCCTAGCTGGGAGGTGATGGCCRGGTCC	1260
Db	1201	GGAGGGTTCGTGACTACATATTAGGCTGTATGAGCCTAGCTGGGAGGTGATGGCCRGGTCC	1260
Qy	1261	ACTGAACCCCTGGTTATCCCAAGAAGCCTTTGCAGGCTTCAGGAGCTTGGAGTGGGGAGAG	1320
Db	1261	ACTGAACCCCTGGTTATCCCAAGAAGCCTTTGCAGGCTTCAGGAGCTTGGAGTGGGGAGAG	1320
Qy	1321	GGGCTGACTTCTCCGACCAAGGCCCTCCACCGGCTACCTTGGGTAAAGGGCCTGGAGCAG	1380
Db	1321	GGGCTGACTTCTCCGACCAAGGCCCTCCACCGGCTACCTTGGGTAAAGGGCCTGGAGCAG	1380
Qy	1381	GAAGCAGGGCAAGAAGCCTCTGGAGCAGCCCATACCGCCCTGGGCTGACTCTGCCACTG	1440
Db	1381	GAAGCAGGGCAAGAAGCCTCTGGAGCAGCCCATACCGCCCTGGGCTGACTCTGCCACTG	1440
Qy	1441	GCAGCACAGTCAACACAGCAGGTTCCTACTCACAGCAGAGGGCAAGGCCATCATCAGCTCC	1500
Db	1441	GCAGCACAGTCAACACAGCAGGTTCCTACTCACAGCAGAGGGCAAGGCCATCATCAGCTCC	1500
Qy	1501	CTTTATTAAGGGAAGGTCAGCGGCTCGGTGTCTCTGAGAGTGTCTCTGTCTCTGTG	1560
Db	1501	CTTTATTAAGGGAAGGTCAGCGGCTCGGTGTCTCTGAGAGTGTCTCTGTCTCTGTG	1560
Qy	1561	CCTGGTGGGGTGGGGTGGCCAGGTGTGCCAGAGAGGCCCATTTGGTAGTGAGGCAGGTA	1620
Db	1561	CCTGGTGGGGTGGGGTGGCCAGGTGTGCCAGAGAGGCCCATTTGGTAGTGAGGCAGGTA	1620
Qy	1621	TGGGGCTAGAAGCAGTGGTGGCCCTGGCCGTGATAGTGGCCATCTTCT	1669
Db	1621	TGGGGCTAGAAGCAGTGGTGGCCCTGGCCGTGATAGTGGCCATCTTCT	1669
RESULT 4			
AAH26179			
ID	AAH26179 standard; DNA; 1669 BP.		
XX	AAH26179;		
AC			
XX			
XX	17-SEP-2001 (first entry)		
DT			
XX	Human cytochrome P450 CYP2D6 gene promoter (G mutation at -1584 bp).		
DE			
XX	Cytochrome P450; CYP2D6; promoter; drug metabolism; human;		
KW	diagnosis; therapy; single nucleotide polymorphism; ds.		
KW			
XX	Homo sapiens.		
OS			
XX			
FH	Key Location/Qualifiers		
FT	variation replace(36.G)		
FT	/*tag= a		
ET	/frequency= "20%"		
FT	5'UTR 1532..1619		
FT	/*tag= b		
FT	CDS 1620..1669		
FT	/*tag= c		
FT	/partial		
FT	/note= "5' region of CYP2D6 coding region"		
XX			
PN	WO200155432-A2.		
XX			
XX	02-AUG-2001.		
XX			
XX	30-JAN-2001; 2001WO-EP00954.		
PF			
XX	31-JAN-2000; 2000EP-0101889.		
PR			
XX			
PA	(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.		



QY 1561 CCTGTGGGGTGGGGTGGCCAGGTGTGTCACAGAGAGCCCATTTGGTAGTGAGGCAGGTA 1620  
 Db 1561 CCTGTGGGGTGGGGTGGCCAGGTGTGTCACAGAGAGCCCATTTGGTAGTGAGGCAGGTA 1620  
 QY 1621 TGGGGCTAGAAGCACTGGTGGCCCTGGCCGTGATAGTGCCCATCTTCCT 1669  
 Db 1621 TGGGGCTAGAAGCACTGGTGGCCCTGGCCGTGATAGTGCCCATCTTCCT 1669  
 RESULT 5  
 ID ABQ72364  
 XX ABQ72364 standard; DNA; 6472 BP.  
 AC ABQ72364;  
 XX  
 DT 02-SEP-2002 (first entry)  
 XX  
 DE Human CYP2D6 gene, SEQ ID NO:1 version #2.  
 XX  
 KW Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;  
 KW Chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;  
 KW antiarrhythmic; arrhythmia; adrenoceptor antagonist; hypertension;  
 KW tricyclic antidepressant; procainamide; drug induced lupus syndrome;  
 KW environmentally linked disease; Parkinson's disease; haplotyping;  
 KW genotyping; haplotype; genetic variant; single nucleotide polymorphism;  
 KW SNP; drug screening; drug discovery; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Location/Qualifiers  
 FT variation  
 FT replace (636, A)  
 FT /tag= a  
 FT /label= PS1  
 FT /note= "Novel single nucleotide polymorphism (SNP)"  
 FT variation  
 FT replace (678, C)  
 FT /tag= b  
 FT /label= PS2  
 FT /note= "Novel single nucleotide polymorphism (SNP)"  
 FT variation  
 FT replace (769, C)  
 FT /tag= c  
 FT /label= PS3  
 FT /note= "Novel single nucleotide polymorphism (SNP)"  
 FT variation  
 FT replace (776, G)  
 FT /tag= d  
 FT /label= PS4  
 FT /note= "Novel single nucleotide polymorphism (SNP)"  
 FT variation  
 FT replace (825, A)  
 FT /tag= e  
 FT /label= PS5  
 FT /note= "Novel single nucleotide polymorphism (SNP)"  
 FT variation  
 FT replace (915, C)  
 FT /tag= f  
 FT /label= PS6  
 FT /note= "Novel single nucleotide polymorphism (SNP)"  
 FT CDS  
 FT 1001..5217  
 FT /tag= g  
 FT /product= "CYP2D6"  
 FT exon  
 FT 1001..1180  
 FT /tag= h  
 FT /number= 1  
 FT variation  
 FT replace (1019, A)  
 FT /tag= i  
 FT /label= PS7  
 FT /note= "Novel single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution V7M"  
 FT variation  
 FT replace (1031, A)  
 FT /tag= j  
 FT /label= PS8  
 FT /note= "Known single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution V11M"  
 FT variation  
 FT replace (1100, T)  
 FT /tag= k

FT  
 FT /label= PS9  
 FT /note= "Known single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution P34S"  
 FT intron  
 FT 1181..1883  
 FT /tag= l  
 FT /number= 1  
 FT variation  
 FT replace (1827, C)  
 FT /tag= m  
 FT /label= PS10  
 FT /note= "Novel single nucleotide polymorphism (SNP)"  
 FT variation  
 FT replace (1843, G)  
 FT /tag= n  
 FT /label= PS11  
 FT /note= "Known single nucleotide polymorphism (SNP)"  
 FT exon  
 FT 1884..2055  
 FT /tag= o  
 FT /number= 2  
 FT variation  
 FT replace (1966, A)  
 FT /tag= p  
 FT /label= PS12  
 FT /note= "Novel single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution R88H"  
 FT variation  
 FT replace (1974, A)  
 FT /tag= q  
 FT /label= PS13  
 FT /note= "Known single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution L91M"  
 FT variation  
 FT replace (1984, G)  
 FT /tag= r  
 FT /label= PS14  
 FT /note= "Novel single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution H94R"  
 FT variation  
 FT replace (1997, G)  
 FT /tag= s  
 FT /label= PS15  
 FT /note= "Novel single nucleotide polymorphism (SNP)"  
 FT variation  
 FT replace (2014, C)  
 FT /tag= t  
 FT /label= PS16  
 FT /note= "Novel single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution V104A"  
 FT variation  
 FT replace (2022, T)  
 FT /tag= u  
 FT /label= PS17  
 FT /note= "Novel single nucleotide polymorphism (SNP);  
 FT together with PS18 causes the amino acid  
 FT substitution T107F"  
 FT variation  
 FT replace (2023, T)  
 FT /tag= v  
 FT /label= PS18  
 FT /note= "Novel single nucleotide polymorphism (SNP);  
 FT together with PS17 causes the amino acid  
 FT substitution T107F"  
 FT variation  
 FT replace (2028, G)  
 FT /tag= w  
 FT /label= PS19  
 FT /note= "Novel single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution I109V"  
 FT variation  
 FT replace (2036, C)  
 FT /tag= x  
 FT /label= PS20  
 FT /note= "Novel single nucleotide polymorphism (SNP)"  
 FT variation  
 FT replace (2039, T)  
 FT /tag= y  
 FT /label= PS21  
 FT /note= "Known single nucleotide polymorphism (SNP)"  
 FT intron  
 FT 2056..2605  
 FT /tag= z  
 FT /number= 2  
 FT /cons.splice= (5'site:NO, 3'site:YES)  
 FT variation  
 FT replace (2062, G)  
 FT /tag= aa  
 FT /label= PS22





```
FT /note= "Known single nucleotide polymorphism (SNP);  
FT 2056..2605 given as Y in the specification"  
FT /tag= z  
FT /number= 2  
FT /cons_splice= (5'site:NO, 3'site:YES)  
FT replace (2062, G)  
FT /tag= aa  
FT /label= PS22  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT given as R in the specification"  
FT replace (2067, G)  
FT /tag= ab  
FT /label= PS23  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT given as K in the specification"  
FT replace (2118, T)  
FT /tag= ac  
FT /label= PS24  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT given as Y in the specification"  
FT replace (2170, A)  
FT /tag= ad  
FT /label= PS25  
FT /note= "Known single nucleotide polymorphism (SNP);  
FT given as R in the specification"  
FT replace (2179, C)  
FT /tag= ae  
FT /label= PS26  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT given as S in the specification"  
FT 2606..2758  
FT /tag= af  
FT /number= 3  
FT replace (2611, A)  
FT /tag= ag  
FT /label= PS27  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT given as W in the specification; causes the  
FT amino acid substitution F120I"  
FT replace (2635, C)  
FT /tag= ah  
FT /label= PS28  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT given as Y in the specification; causes the  
FT amino acid substitution W128R"  
FT replace (2659, A)  
FT /tag= ai  
FT /label= PS29  
FT /note= "Novel single nucleotide polymorphism (SNP);  
FT given as R in the specification; together  
FT with P830 causes the amino acid substitution  
FT V136I"  
FT replace (2661, C)  
  
Query Match 62.9%; Score 1056.6; DB 24; Length 6472;  
Best Local Similarity 99.1%; Pred. No. 3.7e-251;  
Matches 1051; Conservative 10; Mismatches 0; Indels 0; Gaps 0;  
  
QY 620 RACCCTCAGCAGCCCGGAGGATGTTCACAGGTGGGGCAAGGGCCCTCCGGCTACC 679  
Db :|||||  
1 GACCCTCAGCAGCCCGGAGGATGTTCACAGGTGGGGCAAGGGCCCTCCGGCTACC 60  
  
QY 680 AACTGGGAGCTCTGGGAACAGCCCTCTTGCAACAAGAGCCATAGCCCGCCAGAGCCC 739  
Db :|||||  
61 AACTGGGAGCTCTGGGAACAGCCCTCTTGCAACAAGAGCCATAGCCCGCCAGAGCCC 120  
  
QY 740 AGGAATGTGGGCTGGGAGCAGCCCTCTGGACAGGAGTGTCCTCCATCCAGGAAACCT 799  
Db :|||||  
121 AGGAATGTGGGCTGGGAGCAGCCCTCTGGACAGGAGTGTCCTCCATCCAGGAAACCT 180  
  
QY 800 CCGGCATGCTGGGAGTGGGTACTTGGTCCGGGTCTGTATGTGTGTACACTGGTCT 859  
Db :|||||
```

```
Db 181 CCGCATGGCTGGGAAGTGGGTACTTTGGTGCCGGGCTCTGTATGTGTGTGACTGGTGT 240  
QY 860 GTGTGACAGAGAAATGTGTGTCYCTAAGTGTGAGTGTGAGTGTGTATGTGTGAATATGT 919  
Db :|||||  
241 GTGTGACAGAGAAATGTGTGCCCTAAGTGTGAGTGTGAGTGTGTATGTGTGAATATGT 300  
QY 920 CTTTGTGTGGGTGATTTTCTGCRFTGTAAATCGTGTCCCTGCCAAGTGTGACAAAGTGGAC 979  
Db :|||||  
301 CTTTGTGTGGGTGATTTTCTGCGTGTAAATCGTGTCCCTGCCAAGTGTGACAAAGTGGAC 360  
QY 980 AAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTCTGC 1039  
Db :|||||  
361 AAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTCTGC 420  
QY 1040 ATGTCAAGAGTGAAGTGAAGTGAAGGACAGGCCATGATGCCCCACTCATCATCAGGA 1099  
Db :|||||  
421 ATGTCAAGAGTGAAGTGAAGTGAAGGACAGGCCATGATGCCCCACTCATCATCAGGA 480  
QY 1100 GCTCTAAGGCCCCAGGTAAGTGCACGTACAGATAGAGGTGCTCAAGGTCACTCTGGAGT 1159  
Db :|||||  
481 GCTCTAAGGCCCCAGGTAAGTGCACGTACAGATAGAGGTGCTCAAGGTCACTCTGGAGT 540  
QY 1160 GGCAGGTGGGGTAGGAAAGGCAAGGCCATGTTCTGGAGGAGGGGTTGTGACTACAT 1219  
Db :|||||  
541 GGCAGGTGGGGTAGGAAAGGCAAGGCCATGTTCTGGAGGAGGGGTTGTGACTACAT 600  
QY 1220 TAGGTGTATGAGCCTAGCTGGGAGTGGATGGCCRGGTCCACTGAAACCTCGTTATCC 1279  
Db :|||||  
601 TAGGTGTATGAGCCTAGCTGGGAGTGGATGGCCRGGTCCACTGAAACCTCGTTATCC 660  
QY 1280 CAGAAGCTTTGCAGGCTTCAGGAGCTTGGAGTGGGAGAGGGGTGACTCTCCGACCA 1339  
Db :|||||  
661 CAGAAGCTTTGCAGGCTTCAGGAGCTTGGAGTGGGAGAGGGGTGACTCTCCGACCA 720  
QY 1340 GGCCCTCCACCGCCCTACCTGGGTAAGGCCCTGGAGCAGGAGGGGGCAAGAACCT 1399  
Db :|||||  
721 GGCCCTCCACCGCCCTACCTGGGTAAGGCCCTGGAGCAGGAGGGGGCAAGAACCT 780  
QY 1400 CTGGAGCAGCCCATACCCGCCCTGGCTGACTCTGCCACTGGCAGCAGTCAACACAGC 1459  
Db :|||||  
781 CTGGAGCAGCCCATACCCGCCCTGGCTGACTCTGCCACTGGCAGCAGTCAACACAGC 840  
QY 1460 AGTTTCACTCACAGCAGAGGGCAAGGCCATCATCAGTCCCTTTTAAAGGGAGGGTCA 1519  
Db :|||||  
841 AGTTTCACTCACAGCAGAGGGCAAGGCCATCATCAGTCCCTTTTAAAGGGAGGGTCA 900  
QY 1520 CGCGCTCGGTGTGCTGAGAGTGTCTGCGCTGCTCTGTGCTGTGGGTGGGGGTGC 1579  
Db :|||||  
901 CGCGCTCGGTGTGCTGAGAGTGTCTGCGCTGCTCTGTGCTGTGGGTGGGGGTGC 960  
QY 1580 CAGGTGTGCCAGAGAGGCCCATTTGGTAGGAGCAGGTATGGGCTAGAAGCAGTGT 1639  
Db :|||||  
961 CAGGTGTGCCAGAGAGGCCCATTTGGTAGGAGCAGGTATGGGCTAGAAGCAGTGT 1020  
QY 1640 GGCCCTGGCGGTGATAGTGGCCATCTTCCTGCTCTCTGGTGG 1680  
Db :|||||  
1021 GGCCCTGGCGGTGATAGTGGCCATCTTCCTGCTCTCTGGTGG 1061  
  
RESULT 7  
ABK39958  
ID ABK39958 standard; DNA; 5884 BP.  
XX  
AC ABK39958;  
XX  
DT 21-MAY-2002 (first entry)  
XX  
DE Human chemically pretreated gene sequence #20 strand 1.  
XX  
KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;  
KW cytosine; AHD6; CYP11A; CYP11B; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;  
XX UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.  
XX
```















```
CC useful for treating one of the above conditions. The present
CC sequence represents a gene differentially expressed in granulocytes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 122888 BP; 28761 A; 33410 C; 31919 G; 28798 T; 0 other;
PF
XX
Query Match 15.4%; Score 258.2; DB 24; Length 122888;
Best Local Similarity 72.1%; Pred. No. 2.8e-53;
Matches 419; Conservative 3; Mismatches 116; Indels 43; Gaps 5;
Qy 1 GAATTCAGACCCCTGGACAACTTGAAGAACCCSGTCTCTACAAAAATACAAAATT 60
Db 120771 GAATTTGAGACCCCTGGCCAAACATGGTGAACCCCGTCTCTACTAAAAAATACAAA 120830
Qy 61 AGCTGGGATTTGGGTGGCTGCATCGCTATATATATCCAGACACTTTGGGAGCCCTGAGGTG 120
Db 120831 ATTAGGCC--GGGTGGTGGCTGCATCGCTATATATCCAGACACTTTGGGAGCCCTGAGGTG 120888
Qy 121 GGTGATCACCCTGAAGTCAGGAGTTCAAGACTAGCCTGGCCAAACATGGTGAACCCCTATC 180
Db 120889 GCGGATCACCCTGAGTTCGGAGTTTCAGACACCCTTGACCAACATGGAGAAACCCCTGTC 120948
Qy 181 TCTACTGAAATAYAAAA--GCTAGACGTGGTGGCACACACCTGTAAATCCAGCTACT 237
Db 120949 TCTACTAAAAATACAAAATTAGCCGGCGTGGTGGTGCATGCTGTAAATCCAGCTACT 121008
Qy 238 TAGGAGCTGAGGCAGAGAAATTCCTTGAAGCCCTAGAGGTGAAGTTGTAGTGAGCCGAG 297
Db 121009 CAGGAGCTGAGGCAGAGAAATAGCTTGAACCCAGGGAGTGGAGGTGTGGTGAGCCAG 121068
Qy 298 ATTGCATCATTCGACATGAGGAGGGAGCCACCAGCCTGGGCAACAGAGGAAATCTCCGT 357
Db 121069 ATGCGCCATTGCA-----CTCCAGACTGGGCACACAGAGCGAAACACTCTGT 121114
Qy 358 CTCAAAAAATAAAAAAATAAAAAAAGAGT-----TAGGCTGGGTGGTGCCTGTGA 407
Db 121115 TTCCAAAAAATAAAAAAATAAAAAAAGCAAAATTAGCGGGTGTGGTGGCGCCTGTA 121174
Qy 408 GTCCAGCTACTTTGGGAGGCAGGG-----GTCCACTTGATGTCGAGACTG 453
Db 121175 GTCCAGCTACTTTGGGAGGCAGGCGAGGAAGATCGCTTGAACCCAGGAGGAGAGGTTG 121234
Qy 454 CAGTGAGCCATGATCCTCGCACTGCACCTCGCGCTGGGCAACAGAGTGAGACCCCTGTCTA 513
Db 121235 CAGTGAGCTGAGATCGCACAACTACACTCCAGCTCGGACAGAGAGACTCTGTCTCAA 121294
Qy 514 AAGAAAAAATAAAGCAACATATCTCTGAACAAAGGATC 554
Db 121295 AAAAAAATAAAAAAATAAAAAATCCACGAGAAAGGCTC 121335
RESULT 12
AAK78275
ID AAK78275 standard; DNA; 35959 BP.
XX
AC AAK78275;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:33087.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
17-JAN-2001; 2001WO-US01354.
31-JAN-2000; 2000US-0179065.
04-FEB-2000; 2000US-0180628.
24-FEB-2000; 2000US-0184664.
02-MAR-2000; 2000US-0186350.
16-MAR-2000; 2000US-0189874.
17-MAR-2000; 2000US-0190076.
18-APR-2000; 2000US-0198123.
19-MAY-2000; 2000US-0205515.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216647.
11-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
11-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218290.
26-JUL-2000; 2000US-0220963.
26-JUL-2000; 2000US-0220964.
14-AUG-2000; 2000US-0224518.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0225213.
14-AUG-2000; 2000US-0225214.
14-AUG-2000; 2000US-0225266.
14-AUG-2000; 2000US-0225267.
14-AUG-2000; 2000US-0225268.
14-AUG-2000; 2000US-0225270.
14-AUG-2000; 2000US-0225447.
14-AUG-2000; 2000US-0225757.
14-AUG-2000; 2000US-0225758.
14-AUG-2000; 2000US-0225759.
18-AUG-2000; 2000US-0226279.
22-AUG-2000; 2000US-0226681.
22-AUG-2000; 2000US-0227182.
23-AUG-2000; 2000US-0227009.
30-AUG-2000; 2000US-0228924.
01-SEP-2000; 2000US-0229287.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229344.
01-SEP-2000; 2000US-0229345.
05-SEP-2000; 2000US-0229509.
05-SEP-2000; 2000US-0229513.
06-SEP-2000; 2000US-0230437.
06-SEP-2000; 2000US-0230438.
08-SEP-2000; 2000US-0231242.
08-SEP-2000; 2000US-0231243.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231413.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0232080.
08-SEP-2000; 2000US-0232081.
12-SEP-2000; 2000US-0231968.
14-SEP-2000; 2000US-0232397.
14-SEP-2000; 2000US-0232398.
14-SEP-2000; 2000US-0232399.
14-SEP-2000; 2000US-0232400.
14-SEP-2000; 2000US-0232401.
14-SEP-2000; 2000US-0233063.
14-SEP-2000; 2000US-0233064.
21-SEP-2000; 2000US-0233065.
21-SEP-2000; 2000US-0234223.
21-SEP-2000; 2000US-0234274.
25-SEP-2000; 2000US-0234997.
25-SEP-2000; 2000US-0234998.
26-SEP-2000; 2000US-0235484.
27-SEP-2000; 2000US-0235834.
27-SEP-2000; 2000US-0235836.
29-SEP-2000; 2000US-0236327.
29-SEP-2000; 2000US-0236367.
29-SEP-2000; 2000US-0236368.
29-SEP-2000; 2000US-0236369.
```





PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-502630/55.  
XX  
DR Polynucleotides encoding digestive system antigens, useful for  
PT diagnosing, treating, preventing and/or prognosing disorders of the  
PT digestive system, particularly cancer and cancer metastases -  
XX  
PS Disclosure; SEQ ID NO 3234; 986pp; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human digestive system antigens. These can be used in the  
CC diagnosis, treatment and prevention of digestive system disorders,  
CC including cancer, Meckel's diverticulum, bacterial or parasitic  
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or  
CC ulcerative colitis. The present sequence is a genomic DNA fragment  
CC encoding a digestive system antigen of the invention.  
XX  
SQ Sequence 13273 BP; 3179 A; 3261 C; 3701 G; 3132 T; 0 other;

Query Match 15.0%; Score 252.6; DB 22; Length 13273;  
Best Local Similarity 74.0%; Pred. No. 3.2e-52;  
Matches 415; Conservative 3; Mismatches 97; Indels 46; Gaps 6;  
QY 1 GAATTCAGACCGCTGGACAACCTTGGAGAAGCCSGGTCTCTACAAAAATACAAAATT 60

Db 4312 GAGTTCAAAACAGCCTGCCAACATGATGAAACCCCTCTCTACTAAAAATACAAAAA 4371  
QY 61 AGCTGGGATTGGGTGCGGTGCTCATGCTATATATCCAGCACCTTTGGAGCCCTGAGGTG 120  
Db 4372 AGGCC-----GGCGTAGTGGCTCAGGCTGTATATCCCAACACACTTTGGAGGCCAAGGTG 4426  
QY 121 GGTGATCACCTGAAG-ATCAGAGATTCAAGACTAGCCTTGGCCAAACATGGTGAACCCCTAT 179  
Db 4427 GGTGATCACCTGAAGGTTCAGAGTTCAAGACCCAGCCTGGCCACATGGTGAACCTCAT 4486  
QY 180 CTCTACTCAAAATAYAAA----AGCTAGAGCTGGTGGCCACACACCTGTATATCCAGCTA 235  
Db 4487 CTCTACTCAAAATATAAAAAATTAGCCAGGTGTGGTGGCAGGTGCCTGTATATCGTAGCTA 4546  
QY 236 CTTAGAGGCTTGAGGACGAGAATTGCTTGAAGCCTAGAGGTGAAGTGTGAGTGAGCCG 295  
Db 4547 CTCGGGAGCGGAGGTGGGAGAAATCGCTTGAACCTGGAGGTGGAGGTTCAGTGAGCCG 4606  
QY 296 AGATTGCATCATGTCACAAATGGAGGGAGCCAGCCTGGGCAACAGAGAGAAATCTCC 355  
Db 4607 AGATCACCCCATTTGCA-----CTCCAGCCTGGGCAACAGAGCGGAACCTTC 4652  
QY 356 GTCTCCAAAAAATAAAAAAATAAAGRATTAG-----GCTGGGTGGTGCCTGTA 407  
Db 4653 TTCTCAAAAAAATAAAAAAATAAAGRATTAG-----GCTGGGTGGTGCCTGTA 4712  
QY 408 GTCCAGACTACTTGGGAGCGAGGG-----GTCCACTTGATGTCGAGACTG 453  
Db 4713 ATCCAGACTACTCGGAGACTGAGGCATGAAATGGCTTGAACCCGGAGGTGAGGTG 4772  
QY 454 CAGTGAGCCATGATCTTGCACCTGCACTCGGCGTGGGCAACAGAGTGAGACCCCTGCTA 513  
Db 4773 CAGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCGAGACTCTGTCTC 4832  
QY 514 AAGAAAAAATAAAGCAA 534  
Db 4833 AAGAAAAAATAAAGCAA 4853  
RESULT 15  
ID ABA14358 standard; DNA; 32169 BP.  
XX ABA14358;  
AC ABA14358;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Human nervous system related polynucleotide SEQ ID NO 6689.  
XX  
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antiskinking; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200159063-A2.  
XX  
PD 16-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-0501334.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0225759.  
PR 22-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.







GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:51:09 ; Search time 4846.73 Seconds  
(without alignments)  
5613.764 Million cell updates/sec

Title: US-09-942-310-2  
Perfect score: 1680  
Sequence: 1 gaattcaagaccagctgga.....catctctctgctctgtgg 1680

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
EST:*			
1:	em_estba:*		
2:	em_esthum:*		
3:	em_estin:*		
4:	em_estnu:*		
5:	em_estov:*		
6:	em_estpl:*		
7:	em_estro:*		
8:	em_htc:*		
9:	gb_est1:*		
10:	gb_est2:*		
11:	gb_htc:*		
12:	gb_est3:*		
13:	gb_est4:*		
14:	gb_est5:*		
15:	em_estfun:*		
16:	em_estom:*		
17:	gb_gss:*		
18:	em_gss_hum:*		
19:	em_gss_inv:*		
20:	em_gss_pln:*		
21:	em_gss_vrt:*		
22:	em_gss_fun:*		
23:	em_gss_mam:*		
24:	em_gss_mus:*		
25:	em_gss_other:*		
26:	em_gss_pro:*		
27:	em_gss_rod:*		

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	462.4	27.5	1180	14	BQ067695
2	436.6	26.0	449	17	AQ424894
3	414.8	24.7	886	13	BT910844
4	352	21.0	364	17	AQ428360
5	322.6	19.2	382	14	T06700
6	245.6	14.6	631	17	AQ061135

C	7	245.6	14.6	736	17	AQ897828	AQ897828 HS_3135_A
C	8	241	14.3	669	17	AQ635492	AQ635492 RPT-11-4
C	9	235.4	14.0	577	10	AV732057	AV732057 AV732057
C	10	234.6	14.0	661	10	AV730440	AV730440 AV730440
C	11	234	13.9	812	12	BE792565	BE792565 601585351
C	12	231.8	13.8	660	17	AQ057239	AQ057239 CIT-HSP-2
C	13	227	13.5	646	9	AL044489	AL044489 DRF2p4341
C	14	225.8	13.4	738	17	AQ897694	AQ897694 HS_3135_A
C	15	225.6	13.4	593	10	AV762975	AV762975 AV762975
C	16	225	13.4	662	17	AG019380	AG019380 Homo sapi
C	17	225	13.4	664	17	AG019426	AG019426 Homo sapi
C	18	223.4	13.3	539	9	AU158859	AU158859 AU158859
C	19	223.2	13.3	848	17	AQ878774	AQ878774 HS_3135_A
C	20	220.4	13.1	725	17	AG031448	AG031448 Pan trogl
C	21	219.2	13.0	502	14	BQ028890	BQ028890 UI-H-DF0-
C	22	218.8	13.0	529	17	AQ386729	AQ386729 RPT11-14
C	23	217.4	12.9	706	10	AV699423	AV699423 AV699423
C	24	215	12.8	275	12	BF376376	BF376376 MRI-TN004
C	25	214.8	12.8	552	17	AQ599255	AQ599255 HS_3361_A
C	26	214.8	12.8	671	17	AG094205	AG094205 Pan trogl
C	27	214.6	12.8	708	10	AV700654	AV700654 AV700654
C	28	214.4	12.8	710	13	BI603241	BI603241 603249892
C	29	214.2	12.8	647	17	AG065953	AG065953 Pan trogl
C	30	214.2	12.8	658	17	AG094775	AG094775 Pan trogl
C	31	212.8	12.7	522	17	AQ588913	AQ588913 CITBI-EI-
C	32	212.6	12.7	677	12	BG476945	BG476945 602522154
C	33	212.6	12.7	769	17	AQ752554	AQ752554 HS_3572_B
C	34	212.4	12.6	901	12	BG507019	BG507019 601861652
C	35	212.4	12.6	926	14	BQ710345	BQ710345 AGENCOURT
C	36	211.8	12.6	964	14	BQ723625	BQ723625 AGENCOURT
C	37	211.2	12.6	679	17	AG185174	AG185174 Pan trogl
C	38	210.6	12.5	512	17	AQ412029	AQ412029 RPT-11-1
C	39	210.6	12.5	721	9	AL135698	AL135698 DRF2p434J
C	40	210.6	12.5	1034	14	BM917683	BM917683 AGENCOURT
C	41	210.4	12.5	481	13	BM512471	BM512471 1172h07_X
C	42	210.4	12.5	589	17	AQ533097	AQ533097 RPT-11-3
C	43	210.2	12.5	838	12	BG535389	BG535389 602563050
C	44	210	12.5	986	14	BQ710535	BQ710535 AGENCOURT
C	45	209.8	12.5	519	17	AQ427907	AQ427907 CITBI-EI-

ALIGNMENTS

RESULT 1	BQ067695	BQ067695	1180 bp	mrna	linear	EST 02-APR-2002
LOCUS	AGENCOURT_6643098	NIH_MGC_121	Homo sapiens	cdna	clone	IMAGE:5768406
DEFINITION	5', mRNA sequence.					
ACCESSION	BQ067695					
VERSION	BQ067695.1					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 1180)					
AUTHORS	NIH-MGC					
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12828 row: d column: 07 High quality sequence start: 3 High quality sequence stop: 536. Location/Qualifiers					

```
source
1. .1180
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5768406"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."
BASE COUNT      254 a      371 c      347 g      208 t
ORIGIN

Query Match      27.5%; Score 462.4; DB 14; Length 1180;
Best Local Similarity 90.2%; Pred. No. 3.9e-61;
Matches 514; Conservative 3; Mismatches 24; Indels 29; Gaps 1;

QY 575 ATTCTAATCAGAACATGGAGCCACGAAGACGACTGAGGAGGACRACCTCAGGCAGCC 634
Db 29 ATTCTAATCAAAACATGGAGGCCAGGACGACTGAGAAATGACGACCTCAGGCAGCC 88
QY 635 CGGAGGAGTGTGTACAGGCTGGGGCAAGGCCCTCCGGCTACCAACTGGGAGCTCTGG 694
Db 89 CTGGAGGATGCTGTCACAGGCTGGGGCAAGGCCCTTCAGGCTACCAACTGGGAGCTCTGG 148
QY 695 GAACAGCCCTGTTGCCAACAAGAACGACATAGCCGGGCCAGAGCCACGGA----- 743
Db 149 GAACAGCCCTGTTGCCAACAAGAACGACATAGCCGGGCCAGAGCCACGGAATGTGGCTGAG 208
QY 744 -----ATGTGGGCTGGGCTGGGACGACCCCTCTGGACAGGAGTGTGGCTGAG 208
Db 209 CTGGGATCCAGGTGACAGCTTTGAGGCTCACTGGGAGCAGCCCTCTGGACAGGAGTGTCC 268
QY 786 CATCCAGGAAACCTCCGGCATGCTGGGAAGTGGGGTACTTGGTGGCGGGCTGTGTATGTG 845
Db 269 CATCCAGGAAACCTTGGGCATGCTGGGAAGTGGGGTACTTGGTGGCGGGCTGTGTATGTG 328
QY 846 TGTGTGACTGCTGTGTGAGACAGAAATGTGTCYCTAAGTGTCAAGTGTGAGTCTGTGTA 905
Db 329 TGTGTGACTGCTGTGTGAGACAGAAATGTGTCCTGAGTGTCAAGTGTGAGTCTGTGTA 388
QY 906 TGTGTGAATATTCTTGTGTGGGTGATTTCTGCTGTGTGTAATCGTGTCCCTGCAAGT 965
Db 389 TGTGTGAATATTCTTGTGTGGGTGATTTCTGCTGTGTGTAATCGTGTCCCTGCAAGT 448
QY 966 GTGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGGTGTGTG 1025
Db 449 GTGAACAAGTGGACAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGGTGTGTG 508
QY 1026 CATAGGCTCTGTGCATGTCAAGAGTCCAGGTGCAAGTGAAGGACGAGCCCATGATGCC 1085
Db 509 CATAGGCTCTGTGCATGTCAAGAGTCCAGGTGCAAGTGAAGGACGAGCCCATGATGCC 568
QY 1086 ACTCATCATCAGGAGCTCTAAGGCCCCAGG 1115
Db 569 ACTCATCATCAGGAGCTCTAAGGCCCCCG 598

RESULT 2
AQ424894
LOCUS      449 bp      DNA      linear      GSS      24-MAR-1999
DEFINITION      CITBI-E1-2576P11.TF CITBI-E1 Homo sapiens genomic clone 2576P11,
DNA sequence.
ACCESSION      AQ424894
VERSION      AQ424894.1      GI:4498160
KEYWORDS      GSS.
SOURCE      human.
```

```
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 449)
AUTHORS      Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE      Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL      Unpublished (1997)
COMMENT      Other_GSSs: CITBI-E1-2576P11.TR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbeetigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES
source
1. .449
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2576P11"
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"
/Note="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT      134 a      98 c      123 g      94 t
ORIGIN

Query Match      26.0%; Score 436.6; DB 17; Length 449;
Best Local Similarity 97.8%; Pred. No. 5.2e-57;
Matches 439; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 TTCAACACAGCCTGGACAACTTGGGAACACCGGGTCTCTACAAAATATACAAATTAGC 63
Db 1 TTCAAGACAGCCTGTCAAACTTGGGAACACCGGGTCTCTACAAAATATACAAATTAGC 60
QY 64 TGGGATTGGGTGCGGTGCTCATGCTATAATCCAGCACTTTGGGAGCCTGAGGTGGGT 123
Db 61 TGGGATTGGGTGCGGTGCTCATGCTATAATCCAGCACTTTGGGAGCCTGAGGTGGGT 120
QY 124 GGATCACCTGAAAGTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAACCCCTATCTCT 183
Db 121 GGATCACCTGAAAGTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAACCCCTATCTCT 180
QY 184 ACTGAAATATYAAAAGCTAGACGTGGTGGGCACACACCTGTAAATCCAGCTACTTAGGAG 243
Db 181 ACTGAAATATACAAAAGCTAGACGTGGTGGGCACACACCTGTAAATCCAGCTACTTAGGAG 240
QY 244 GCTGAGGAGGAGAATTGCTTGAAGCCTAGAGGTGAAGTTGTTAGTGAAGCCGAGATTGCA 303
Db 241 GCTGAGGAGGAGAATTGCTTGAAGCCTAGAGGTGAAGTTGTTAGTGAAGCCGAGATTGCA 300
QY 304 TCATTGCAATATGGAGGGAGCCACGACCTGGGCAACAAGAGAAATCTCCGTCTCAA 363
Db 301 TCATTGCAATATGGAGGGAGCCACGACCTGGGCAACAAGAGAAATCTCCGTCTCAA 360
QY 364 AAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 423
Db 361 AAAAAAATTAATACAAAAGATTAAGGCTGAGTGGTGCCTGTAGTCCAGCTACTTAGG 420
QY 424 AGGAGGGGGTCCACTTGTGATGTCGAGACT 452
Db 421 AGGAGGGGGTCCACTTGTGATGTCGAGACT 449

RESULT 3
```



```

QY 244 GCTGAGCGAGGAATTCCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCGAGATTGCA 303
|||||
Db 241 GCTGAGCGAGGAATTCCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCGAGATTGCA 300
|||||
QY 304 TCATTGCACATGGAGGAGGACCGACCGCTGGGGCAACAAGAGGAAATCTCCGTCTCAA 363
|||||
Db 301 TCATTGCACATGGAGGAGGACCGACCGCTGGGGCAACAAGAGGAAATCTCCGTCTCAA 360
|||||
QY 364 AAAA 367
|||
Db 361 TATA 364
|||||

RESULT 5
T06700/c
LOCUS
DEFINITION
T06700 382 bp mRNA linear EST 30-JUN-1993
EST04589 Fetal brain, Stratagene (cat#936206) Homo sapiens cDNA
clone HFBDX16 similar to EST containing Alu repeat, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
T06700 Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT
Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.
3,400 expressed sequence tags identify diversity of transcripts
from human brain
Nat. Genet. 4, 256-267 (1993)
93364420
Contact: Adams, MD
The Institute for Genomic Research
932 Clopper Road, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423
Email: mdadams@tigr.org
Seq primer: M13-21.

FEATURES
source
Location/Qualifiers
1..382
/db_xref="ATCC (inhost):83354"
/db_xref="taxon:9606"
/clone="HFBDX16"
/clone_lib="Fetal brain, Stratagene (cat#936206)"
/notes="Vector: LambdaZAP-II; 17-18 wk gestation, female;
oligo-dT + random primed cDNA synthesis; lambdaZAP-II
vector, 1.0kb average inser size."
BASE COUNT 76 a 101 c 90 g 111 t 4 others
ORIGIN
Query Match 19.2%; Score 322.6; DB 14; Length 382;
Best Local Similarity 97.9%; Pred. No. 1.1e-39;
Matches 333; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 22 AACTTGGAGAACCGGTCTCTACAAAATA-CAAAATAGCTGGGATGGGTGGCGTG 80
|||||
Db 340 AACTTGGAAACCGGGTCTCTACAAAATNCCAAATAGCTGGGATGGGTGGCGTG 281
|||||
QY 81 GCTCATGCCTATAATCCCGACACTTTGGGAGCTGAGGTGGGTGATCACCTGAAGTCAG 140
|||||
Db 280 GCTCATGCCTATAATCCCGACACTTTGGGAGCTGAGGTGGGTGATCACCTGAAGTCAG 221
|||||
QY 141 GAGTTCAACACTAGCTGGCCACATGTTGAACCCCTATCTCTACTGAAATAYAAAAAG 200
|||||
Db 220 GAGTTCAACACTAGCTGGCCACATGTTGAACCCCTATCTCTACTGAAATAYAAAAAG 161
|||||
QY 201 CTAGAGCTGGTGGCACACACCTGTATCCCGACTACTTAGGAGGCTGAGCGAGGAGATT 260
|||||
Db 160 CTAGAGCTGGTGGCACACACCTGTATCCCGACTACTTAGGAGGCTGAGCGAGGAGATT 101
|||||
QY 261 GCTTGAAGCCTAGAGGTGAAGTTTGTAGTGAGCGGAGATTGTCATCATTGCACAAATGAGG 320
|||||

```

```

Db 100 GCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCGAGATTGTCATTCGACAAATGAGG 41
|||||
QY 321 GGAGCCACCAGCCTGGGCAACAAGAGGAAATCTCCGTCTC 360
|||||
Db 40 GGAGCCACCAGCCTGGGCAACAAGAGGAAATCTCCGTCTC 1
|||||

RESULT 6
AQ061135/c
LOCUS
DEFINITION
CIT-HSP-2348E15.TF CIT-HSP Homo sapiens genomic clone 2348E15, DNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AQ061135.1 GI:3363047
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Other_GSSs: CIT-HSP-2348E15.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..631
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2348E15"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/notes="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
BASE COUNT 116 a 162 c 147 g 206 t
ORIGIN
Query Match 14.6%; Score 245.6; DB 17; Length 631;
Best Local Similarity 72.4%; Pred. No. 3.9e-28;
Matches 414; Conservative 3; Mismatches 112; Indels 43; Gaps 6;

QY 2 AATTCAAGACACGAGCTGGACACACTTGAAGAACCGGCTCTACAAAAATACAAATTA 61
|||||
Db 604 AGTTCAAGACACGAGCTGGCCACACATGGTGAACCCCATCTCTACTGAAATACCAAAAA 545
|||||
QY 62 GCTGGGATT-----GGGTGCGTGGCTCATGCCTATATATCCAGCACCTTTGGGAGCCTGA 116
|||||
Db 544 AAAAAAAGAGCGGCTCAGTGGCTCACCGCTTCTATCTTAGCACACTTTGGGAGGTGA 485
|||||
QY 117 GGTGGTGGATCACCCTGAAGTCAGAGTTCAAGACTAGCCTGGCCAAACATGGTGAACACC 176
|||||
Db 484 GACGGGCGGTTCATCTGAGGTCGGGAGTTCAAGACCGCCTGACCATCATGTTGAACACC 425
|||||
QY 177 TATCTCTACTGAAATAYAAAA--AGCTAGACGTGGTGGCACACACCTGTATATCCAGCT 234
|||||
Db 424 -GTCTCTACTAATAATACAAATTTACCGGGCATAGTGGCATATGCCTCTAATATCCAGCT 366
|||||

```

Qy	235	ACTTAGGAGCGTGGAGGAGAATTCCTTGAAACCCCTAGAGGTGAAGGTTGTAGTGAGCC	294
Db	365	ACTGGGAGCGTGGAGGAGGAGAATTCCTTGAAACCCAGGAGCGGAGGTTGCATGAGCC	306
Qy	295	GAGATTGCATCATGACAATGGAGGGAGGCCACCAGCCTGGGCAACAAGAGAAATCTC	354
Db	305	ANGATCACCAATTGCA-----CCTTCAGCCTGGGCAACAAGAGTGAACACTC	260
Qy	355	CGTCTCCAATAAAAAAAAAAAAAAAGATTTAGGCTGGGTGG-----TGCCCTGTA	407
Db	259	CATCTCAAAAAAAAAAAAAAATAATAGCGGCGCTGGTGGCTGCCCTGCCCTGTA	200
Qy	408	GTCCAGTACTTGGGAGGCGAGGG-----GTCCACTTGTAGTGTGAGACTG	453
Db	199	GTCCAGCTATTTGGGAGGCTGAGCATGATGAGAACTCGTTGAACCCAGGAGCAGAGTTG	140
Qy	454	CAGTGAGCCATGATCCTGCCACTGCACCTCCGGCCTGGGCAACAGCTAGACCCCTGCTA	513
Db	139	CAGTGAGCCAGATGACATGACTGCATCCAGCTGGGCAACAGAGTTAGACTCTGTCCC	80
Qy	514	ANGAAAAAAAATAAAGCAACATATCTCTGAA	545
Db	79	AAAAAAGAAAAAATAAAGCTTGA	48
RESULT 7	AQ897828/c		
LOCUS	HS_3135_A2_All_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-3135 Col=22 Row=A, DNA sequence.	736 bp	linear GSS 10-NOV-1999
ACCESSION	AQ897828		
VERSION	AQ897828.1	GI:6353934	
KEYWORDS	GSS.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 736) Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Wood,L.		
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)		
MEDLINE	99380589		
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones may be purchased from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu Plate: 3135 row: A column: 22 Seq primer: T7 Class: BAC ends High quality sequence stop: 736. Location/Qualifiers 1..736 /organism="Homo sapientis" /db_xref="taxon:9606" /clone="plate-3135 Col=22 Row=A" /clone.lib="CIT Approved Human Genomic Sperm Library D" /sex="male" /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"		
BASE COUNT	142 a	196 c	159 g 238 t 1 others
ORIGIN			
Query Match	14.6%	Score 245.6;	DB 17; Length 736;
Best Local Similarity	71.2%;	Pred. No. 3.6e-28;	
Matches 423; Conservative	2;	Mismatches 126;	Indels 43; Gaps 6;





ACCESSION AV730440  
VERSION AV730440.1 GI:10839861  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 661)  
AUTHORS Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,  
Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu  
.S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,  
Chen,J., Chen,Z. and Han,Z.  
Homo sapiens cDNA Htf clones  
Unpublished (2000)  
JOURNAL  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chc.sh.cn  
This clone is available at CHGC in Shanghai.  
FEATURES  
Location/Qualifiers  
source  
1..661  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HTFAWE01"  
/clone\_lib="Htf"  
/tissue\_type="Hypothalamus"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"  
BASE COUNT 183 a 178 c 179 g 116 t 5 others  
ORIGIN  
Query Match 14.0%; Score 234.6; DB 10; Length 661;  
Best Local Similarity 72.7%; Pred. No. 1.8e-26;  
Matches 380; Conservative 3; Mismatches 107; Indels 33; Gaps 5;  
QY 1 GAATTCACACAGCCTGGACACTTGGAGAACCSGGTCTCTACAAAAAATACAAATT 60  
DB 8 GATTTCCAGACAGCCTGACCAACATGGAGAACCCCTGTCTACTAAAAAGTATAGAA 67  
QY 61 AGCTGGGATTGGGTGGGTGCTCATGTCCTATAATCCAGCAGCTTTGGGAGCCTGAGGTG 120  
DB 68 TT----GGCCGGCGGGTGGCTCAGCCCTGTATCCAGCAGCTTTGGGAGCCGAGCG 123  
QY 121 GGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCCTGGCCACACATGGTGAACCCCTATC 180  
DB 124 GGCAGATCACCTGAGGTGGGAGTTCCAGCAGCAGCTGACCAACATGGAGAAACCCCTGTC 183  
QY 181 TCTACTGAAATAYAAAA--ACGTAGAGTGGTGGCAGACACACCTGTAATCCAGCTACT 237  
DB 184 TCTACTGAAATATACAAAAATTAGCCGGGCGTGGTGCATGCCCTATAATCCAGCTACT 243  
QY 238 TAGGAGGCTGAGGCGAGGAATTGCTTTGAAGCCTAGAGGTGAAGGTGTAGTGAGCCGAG 297  
DB 244 TGGTAGCGGAGCAGGAGAAATCGCTTTGAACCCGGGAGCGAGGTTCAGTGAGCCGAG 303  
QY 298 ATTGCATCATTTGCACAAATGGAGGGAGCCAGCAGCTGGGCAACAGAGAAATCTCCGT 357  
DB 304 ATCGGCCCACTGCACCT-----CCAGCTTTGGCAACAGAGCGAAACCTCCAC 349  
QY 358 CTCCAAAAAAANAANAANAANAANGRA-----TTAGGCTGGGTGGTGCCTGTATA 407  
DB 350 CTCAAAAAANAANAANAANAANGAAATTAGCCAGGCATGGTGGCAGCAGCCGTATA 409  
QY 408 GTCCCAAGCTACTTTGGAGGC--AGGGGTCTCACTTGTATCGAGACTTCAGGTGAGCCATG 465  
DB 410 GTCCCACTACTCTACGAGAGACTGAGGCAGAACCCAGGAGGTGGAGACTGCGATCAGCCAA 469  
QY 466 ATCCTGCCACTGCACCTCCCGCCTGGGCAACAGAGTGAAGACCT 508

DB 470 ATGACACCAGCTGCACCTCCAGCCTGGGAGGAAGAGTGAGACTCT 512  
RESULT 11  
BE792565 812 bp mRNA linear EST 20-SEP-2000  
LOCUS 601585351F1 NIH\_MGC\_7 Homo sapiens cDNA clone IMAGE:3939468 5',  
DEFINITION mRNA sequence.  
ACCESSION BE792565  
VERSION BE792565  
KEYWORDS BE792565.1 GI:10213763  
SOURCE EST.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 812)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: DCTD/Drp  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing: The I.M.A.G.E. Consortium (LLNL)  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LLCW788 row: f column: 13  
High quality sequence stop: 683.  
Location/Qualifiers  
source  
1..812  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3939468"  
/clone\_lib="NIH\_MGC\_7"  
/tissue\_type="small cell carcinoma"  
/cell\_line="MGC3"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pOT87; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGACAGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life technologies)."  
BASE COUNT 215 a 177 c 247 g 173 t  
ORIGIN  
Query Match 13.9%; Score 234; DB 12; Length 812;  
Best Local Similarity 70.9%; Pred. No. 2e-26;  
Matches 393; Conservative 2; Mismatches 122; Indels 37; Gaps 5;  
QY 1 GAATTCAGACACAGCCTGGACAACTTTGGAGAACCSGGTCTCTACAAAAAATACAAATT 60  
DB 152 GAGTTTGGAGACAGCCTGGGCAACATGGTGAAGTCCATCTCTACAAAAAATACAAATT 211  
QY 61 AGCTGGGATTGGGTGGGTGGGTGCTCATGTCCTATAATCCAGCAGCTTTGGGAGCCTGAGGTG 120  
DB 212 CA----GGTGGGTGGGTGAGTGCACACCTCTAATCCAGCAATTTGGGAGGTGAAGTA 267  
QY 121 GGTGGATCACCTCAAGTCAGGAGTTCAAGACTAGCCTGGCCACACATGGTGAACCCCTATC 180  
DB 268 AGTGGATCACTGAGGTGAGGTTCAGAGCTTCAAGCAGCAGCTGGCCAAACATGGCCAAACCCCGTC 327  
QY 181 TCTACTGAAATAYAAAA--AGCTAGAGTGGTGGCAGACACACCTGTAATCCAGCTACT 237  
DB 328 TCTGTCTGAAATAGAAAAATCAGCTGGCATGGTGGCGTGTGCTGTGTAATCCAGCTACT 387  
QY 238 TAGGAGGCTGAGGCGAGGAATTCGTTGAAGCCTAGAGGTGAAGGTGTGTAAGTGAAGCCGAG 297  
DB 388 CAGGAGGCTGAGGCTGAGAAATCGCTTGAACCCGGGAGCGGAGGTGCGAGTGAAGTGAAG 447



BASE COUNT	137 a	168 c	136 g	205 t	
ORIGIN	/clone="DKF2P434I202" /clone_lib="434 (synonym: htes3)" /tissue_type="testis" /dev_stage="adult" /lab_host="DH10B" /notes:Vector: pSport1; Site_1: NotI; Site_2: SalI"				
Query Match	13.5%	Score 227;	DB 9;	Length 646;	
Best Local Similarity	70.6%	Pred. No. 2.5e-25;			
Matches 385; Conservative	3;	Mismatches 118;	Indels	39;	Gaps
QY 1	GAATTCAGACACCGCTGGACAAC	TTGGAAGAACCCGGTCTCTACAA	AAAAATACAAAATT	60	
Db 526	GAGTTTGAGACACGACATGGCAAC	ATGGTGAACCCCGCTCTACAAAAT	ACTAAATA	467	
QY 61	AGCTGGGATTGGGTGGCGTGGCT	CATCGCTATATATCCAGCACATT	TTGGGAGCCCTGAGGTG	120	
Db 466	AGGCAAGAAG	-----GGTGGCTCATGCCATA	TATCCAGCATTTTGGGAGGCCAAGCG	412	
QY 121	GGTGATCACCTGAAGTCAGGAGT	TCAAGACTAGCCATGGCCACACAT	GGTGAACCCCTATC	180	
Db 411	GGCAGATCACTTGAAGTCAGGAG	TTTGAGACACGATGGCCACACAT	GATGAACCCCGCTC	352	
QY 181	TCTACTGAAATAYAAAA	---AGCTAGACGCTGGTGCCACAC	CACCTGTAAATCCCAGCTACT	237	
Db 351	TCTACTAAAAATATAAAAAT	TAGCTGGCGATGGTGGTGGCCAC	CTGTAAATCCCAGCTACT	292	
QY 238	TAGGAGCTGAGGCAGAGAAATTC	CTTTGAAGCCCTAGAGGTGAAG	GTGTAGTGAGCCGAG	297	
Db 291	CTGGAGCTGAGGCAGAGAAATCA	TTTGAACCTTAGGAGGCAGAG	GTGTGCATGAGCCCAAG	232	
QY 298	ATTGCATCTATTGCAACATGG	AGGGGAGCCACCGCTGGGCCAAC	AGAGGAAATCTCCGT	357	
Db 231	ATTGCGTCACTGCA	-----TTCCAGCCCTGGGCCAAC	AGAGCAAGCAACTCCAT	186	
QY 358	CTCCAAAAA	AAAAA	AAAAA	AGAGTAT	414
Db 185	CTAAAAA	AAAAA	AAAAA	ATAGCAGGATGTGGTGGTGTCTCTCTAGTCTCAG	126
QY 415	CTACTTGGGAGGCAGGG	-----GTCCACTTGATTCGAGACTGCAGTGAG	460		
Db 125	CTACTCAGAGGCTTGAGGCAG	AAGAAATCACTTGAACCCAGGAGTGGAGATTCGAGTTAA	66		
QY 461	CCATGATCTCTGCCACTGCAC	TCCGGCTGGGCCAACAGACTGAGACCCCTGTCTAAAGAAA	520		
Db 65	CTGAGATAGGCCACTGCATTC	ACGCTTCCAGCTTGGATGACAGTGAGACTCCATCTTAAAAAA	6		
QY 521	AAAAA	525			
Db 5	AAAAA	1			
RESULT 14	A0897694/c				
LOCUS	A0897694				
DEFINITION	HS_3135.A2.B12.T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-3135 Col-24 Row-C, DNA sequence.				
ACCESSION	A0897694				
VERSION	A0897694.1				
KEYWORDS	GSS.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	1 (bases 1 to 738) Mahalax, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.				
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)				

```

99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com)
BAC end Web Server: http://www.htsc.washington.edu
Plate: 3135 row: C column: 24
Seq primer: T7
Class: BAC ends
High quality sequence stop: 738.
FEATURES
    source                1..738
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="Plate-3135 Col-24 Row-C"
                        /clone_lib="CIT Approved Human Genomic Sperm Library D"
                        /sex="male"
                        /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones In
                                E-Coli DH10B"
BASE COUNT      147 a   197 c   152 g   241 t       1 others
ORIGIN
Query Match      13.4%; Score 225.8; DB 17; Length 738;
Best Local Similarity 69.0%; Pred. No. 3.6e-25;
Matches 412; Conservative 1; Mismatches 143; Indels 41; Gaps

Qy  3  ATTCAAGACCAGCCTGGACAACTTTGGAGAACCSGGTCTCTACAAAAAATACAAAAATTAG 62
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  603 ACTAGAGACCAGCCTGGCAACATAGCAAGACCTCATCTTTACAAATAAATTTAAA---AA 547

Qy  63  CTGGGATGGGTGGGTGGCTCATGTCCTATATATCCAGCACTTTTGGGAGCCTGAGGTGGG 122
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  546 ATTAGGCTGGGTGGGTGGCCACACTGTATATCCAGCACTTTTGGGAGGCGCAAGCAGG 487

Qy  123 TGGATCACTGAAGTCAGGAGTTCAAGACTAGCCTGGCCAACTGGTGAACCCCTATCTC 182
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  486 CGATCACTGAGTCAGGAGTTTGAGACCAGCCTGACCACATGGAGAAACCCCATCTT 427

Qy  183 TACTGAAATAATYAAAA--AGCTAGACGTGGTGGCACACACCTGTAAATCCCAAGCTACTTAG 240
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  426 TACTAAAAATACAAAATTAGCCTGCTGGTGGGCACATGCCTGTAAGCCAGCTACTCAG 367

Qy  241 GAGCCTGAGGCAGAGAAATTCCTTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCCGAGATT 300
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  366 GAGGCTGAGGCAGAGAAATTAATCTGAACCTGGTAGGGCGAGTTTGTGGTGAAGCTGAGTT 307

Qy  301 GCATCTTGCAATGGAGGGGAGCCAGCCTGGCAACAGAGGAAATCTCCGCTCTC 360
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  306 GCGCCATTGCA-----CTTCAGCTTTGGCAACAGGGGCAAAAATTCGTTTT 261

Qy  361 CAAAAAATAAAAAAAAAAAAAAGRA-----TTAGCTGGGTGGTGCCTGTAGTCCCA- 413
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  260 CAAAAAATAAAAAAAAAAAAAAATTTAGCCAGGGGTGGGTGSCAAAAAACCTGTATTCAG 201

Qy  414 -----GCTACTTTGGAGGCGAGGGGTCCACTTGTATGTCGAGACTGCAAGTGCAGTGCAGC- 461
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  200 TTATCAGAGGCTGAGTGGGAGGATCTCGAGCCCAAGAGGTGCGAGGTGCGAGTGCAGTGT 141

Qy  462 --CATGATCTCGCACTGCACCTCCGGCCTGGGCAACAGACTGAGACCCCTGTCTAAAGMA 518
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  140 GATTGTGATCTTGGCACTGCACTTCAGCGTGGGAGACAGCAAGACCCCACTCATTTAA 81

Qy  519 AAAAAAATAAAGCAACATATCCTTGAACAAAGGATCCTCCATAACGTTTCCACCACA 575
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  80 AAAAAAATAAATAAAGAGAGAGAGAGAGAGAGAGATACAGCAGCTTCCCAAGTAA 24

RESULT 15
AV762975

```

LOCUS AV762975 593 bp mRNA linear EST 19-OCT-2000  
DEFINITION AV762975 MDS Homo sapiens cDNA clone MDSCK10 5', mRNA sequence.  
ACCESSION AV762975  
VERSION AV762975.1 GI:10920823  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 593)  
Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,  
Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng  
L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,  
Yang,Y., Gao,G., Zhang,Q., Chen,S., Han,Z. and Chen,Z.  
Homo sapiens cDNA MDS clones  
Unpublished (2000)  
Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.  
FEATURES  
source  
1..593  
Location/Qualifiers  
1..593  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MDSCK10"  
/clone\_lib="MDS"  
/tissue\_type="Bone marrow"  
/cell\_type="CD34+ hematopoietic stem/progenitor cell"  
/lab\_host="BM25.8"  
/note="Vector: pTriplex2; Site\_1: sf1A; Site\_2: sf1B"  
BASE COUNT 169 a 143 c 170 g 110 t 1 others  
ORIGIN  
Query Match 13.4%; Score 225.6; DB 10; Length 593;  
Best Local Similarity 71.0%; Pred. No. 4.3e-25;  
Matches 379; Conservative 3; Mismatches 117; Indels 35; Gaps 5;  
QY 1 GAATTCAGACACCGCTGGACAACTTGGAGAAGCCSGGTCTACAAAAAATACAAAAT 60  
Db 66 GAGTTCAATGCCAGCCCGAGCAACATAGTGAGACCCCTGTCTATTATTTAAAGAAATAGTA 125  
QY 61 AGCTGGGATTGGGTGGCGTGCCTCATGCCCTAATATCCAGCACTTTGGAGCCTGAGGTG 120  
Db 126 ATAAAGSGCTGGGCACGGTGGCTTACGCCCTGTAATCCAGCACTTTGGAGGCGCGAGGTG 185  
QY 121 GGTGGATCAGCTGAAGTCAGAGTTCAAGACTAGCCTGGCCAAACATGGTCAAAACCCCTATC 180  
Db 186 GGTGGATCAGCTGAGGTGAGGATCAAGACCAGCCCTAGCCAAACATGGTCAAAACCCCGTC 245  
QY 181 TCTACTGAAAAATAVAAAA-----AGCTAGACGTGGTGGCACACACCTGTGAATCCAGCTA 235  
Db 246 TCTACTAAAAAAACAAAAATTTAGCCGGCGATGGTGGCATGTGCCTGTAAATCCCACTA 305  
QY 236 CTTAGGAGGCTGAGGCAGGAGAAATTCCTTGAAGCCTTAGAGGTGAAGGTTGTAGTGAGCGG 295  
Db 306 CTCGGGAGGCTGTGGCAGAAGAATTGCTTGAACCCAGGAGGAGATGTTGCAGTGAGCGG 365  
QY 296 AGATTGCATCATTTGCACATGGAGGGAGGCCACCGCTGGGCAACAAAGGAAATCTCC 355  
Db 366 AGATCGTGCACCTGCACACT-----CCAGCCTGGGCAA-AAGAGCAAGAGTCC 410  
QY 356 GTCTCAAAAAAAGAAAAAAGAAAAAGAAATTAAGCTGGTGGTGCCTGTAGTCCAGC 415  
Db 411 ATCTCAAAAAAAGAAAAAAGAAAAATAGCCAGGCGTGG-TGGCAGGACCTGTATATCCAGC 469  
QY 416 TACTTGGGAGGCGAGGG-----GTCCACATTGATGCGAGACTGCAGTGAGC 461  
Db 470 TACTTGGGAGGCTGAGGAGGAGAAATCGCCTGAACCTAGGAGGCAGAGGTTCCAGTAAGC 529

QY 462 CATGATCCTGCCACTCCACTCCGGCCTGGGCAACAGAGTGAGACCCCTGTCTAAA 515  
|| ||| ||||||||||||| ||||||||||||||||||||| ||| |||  
Db 530 CAAGAATTGCGCCACTGCACCTCCAGCCTGGGCAACAGAGTGAGACTCTGCTTCAA 583

Search completed: February 11, 2003, 03:31:02  
Job time : 4853.73 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 21:12:29 : Search time 127.662 Seconds  
(without alignments)  
4035.781 Million cell updates/sec

Title: US-09-942-310-2

Perfect score: 1680

Sequence: 1 gaattcaagaccagctgga.....catctctctgctctgtgg 1680

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	250	14.9	10684	3	US-08-618-100B-3
2	244.2	14.5	62804	4	US-09-800-960-3
3	233.6	13.9	11531	1	US-08-068-945A-1
4	233.6	13.9	11531	1	US-08-442-806-1
5	232.8	13.9	15297	4	US-09-817-180-3
6	229	13.6	4192	4	US-09-122-126B-1
7	215.4	12.8	246240	2	US-08-724-394A-20
8	215.4	12.8	246240	2	US-08-724-394A-21
9	215.4	12.8	246240	2	US-08-724-394A-22
10	215.2	12.8	6235	4	US-09-305-384-5
11	215.2	12.8	6679	4	US-09-305-384-1
12	215	12.8	11811	4	US-09-078-294-7
13	214.6	12.8	8174	1	US-07-914-281-5
14	214.6	12.8	8174	1	US-08-393-246-5
15	214.6	12.8	8174	1	US-08-525-058A-5
16	214.6	12.8	8174	2	US-08-696-731-5
17	214.6	12.8	8174	4	US-09-042-531-5
18	214.6	12.8	8174	5	PCT-US91-00899-3
19	214.4	12.8	43950	4	US-09-735-934A-3
20	213.8	12.7	162450	4	US-09-345-882-1
21	212.2	12.6	162450	4	US-09-345-882-1
22	211.8	12.6	99500	4	US-09-798-096-10
23	208.8	12.4	59065	4	US-09-813-817-3
24	208.8	12.4	59065	4	US-09-978-197-3
25	207.2	12.3	70000	4	US-09-851-896-3
26	206.2	12.3	112332	4	US-09-741-150-3
27	203.2	12.1	98844	4	US-09-791-211-10
1	250	14.9	10684	3	Sequence 3, Appli
2	244.2	14.5	62804	4	Sequence 3, Appli
3	233.6	13.9	11531	1	Sequence 1, Appli
4	233.6	13.9	11531	1	Sequence 1, Appli
5	232.8	13.9	15297	4	Sequence 3, Appli
6	229	13.6	4192	4	Sequence 1, Appli
7	215.4	12.8	246240	2	Sequence 20, Appli
8	215.4	12.8	246240	2	Sequence 21, Appli
9	215.4	12.8	246240	2	Sequence 22, Appli
10	215.2	12.8	6235	4	Sequence 5, Appli
11	215.2	12.8	6679	4	Sequence 1, Appli
12	215	12.8	11811	4	Sequence 7, Appli
13	214.6	12.8	8174	1	Sequence 5, Appli
14	214.6	12.8	8174	1	Sequence 5, Appli
15	214.6	12.8	8174	1	Sequence 5, Appli
16	214.6	12.8	8174	2	Sequence 5, Appli
17	214.6	12.8	8174	4	Sequence 5, Appli
18	214.6	12.8	8174	5	Sequence 3, Appli
19	214.4	12.8	43950	4	Sequence 3, Appli
20	213.8	12.7	162450	4	Sequence 1, Appli
21	212.2	12.6	162450	4	Sequence 1, Appli
22	211.8	12.6	99500	4	Sequence 10, Appli
23	208.8	12.4	59065	4	Sequence 3, Appli
24	208.8	12.4	59065	4	Sequence 3, Appli
25	207.2	12.3	70000	4	Sequence 3, Appli
26	206.2	12.3	112332	4	Sequence 3, Appli
27	203.2	12.1	98844	4	Sequence 10, Appli

Sequence 8, Appli  
Sequence 8, Appli  
Sequence 8, Appli  
Sequence 42, Appli  
Sequence 59, Appli  
Sequence 10, Appli  
Sequence 11, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 23, Appli  
Sequence 32, Appli  
Sequence 43, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 57, Appli  
Sequence 3, Appli

Sequence 8, Appli  
Sequence 8, Appli  
Sequence 8, Appli  
Sequence 42, Appli  
Sequence 59, Appli  
Sequence 10, Appli  
Sequence 11, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 23, Appli  
Sequence 32, Appli  
Sequence 43, Appli  
Sequence 2, Appli  
Sequence 3, Appli  
Sequence 57, Appli  
Sequence 3, Appli

## ALIGNMENTS

RESULT 1  
US-08-618-100B-3  
: Sequence 3, Application US/08618100B  
: Patent No. 6068976  
: GENERAL INFORMATION:  
: APPLICANT: Briggs, Michael R.  
: APPLICANT: Auwerx, Johan  
: APPLICANT: de Vos, Piet  
: APPLICANT: Staels, Bart  
: APPLICANT: Croston, Glenn E.  
: APPLICANT: Miller, Stephen G.  
: TITLE OF INVENTION: MODULATORS OF ob GENE AND  
: TITLE OF INVENTION: SCREENING METHODS THEREFOR  
: NUMBER OF SEQUENCES: 48  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: Lyon & Lyon  
: STREET: 633 West Fifth Street  
: STREET: Suite 4700  
: CITY: Los Angeles  
: STATE: California  
: COUNTRY: U.S.A.  
: ZIP: 90071-2066  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
: MEDIUM TYPE: Storage  
: COMPUTER: IBM Compatible  
: OPERATING SYSTEM: IBM P.C. DOS 5.0  
: SOFTWARE: FastSeq for Windows Version 2.0  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/618,100B  
: FILING DATE: March 19, 1996  
: CLASSIFICATION: 514  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: 08/558,588  
: FILING DATE: October 30, 1995  
: APPLICATION NUMBER: 08/510,584  
: FILING DATE: August 2, 1995  
: APPLICATION NUMBER: 08/418,096  
: FILING DATE: April 5, 1995  
: APPLICATION NUMBER: 08/408,584  
: FILING DATE: March 20, 1995  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Warburg, Richard J.  
: REGISTRATION NUMBER: 32,327  
: REFERENCE/DOCKET NUMBER: 219/075  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: (213) 489-1600  
: TELEFAX: (213) 955-0440  
: TELEX: 67-3510  
: INFORMATION FOR SEQ ID NO: 3:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10684 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; DESCRIPTION: Sequence between exon 1 and exon 2
; Patent No. 6068976
US-08-618-100B-3
```

```
Query Match 14.98; Score 250; DB 3; Length 10684;
Best Local Similarity 72.18; Pred. No. 4.4e-57;
Matches 404; Conservative 3; Mismatches 118; Indels 35; Gaps 5;

QY 1 GAATCAAGACCAAGCTGGACAACTTGAAGAACCCTGCTCTCTCAAAAAATACAAATT 60
Db 6778 GAGTCAAGACCAAGCTGGCCAAACATGGTGAACCCCATCTCTACTAAAAATACAAAA 6837

QY 61 AGCTGGGATTTGGGTGGCGTGCCTCATGCTATAATCCAGCACTTTGGGAGCCTTAGG 120
Db 6838 A--TGAGGCTGGGCGTGATGACTCACACCTGTAATCCAGCACTTTGGGAGCGCGAG 6895

QY 121 GGTGATCACCTGAAGTCAGGAGTTCAGAGTACGCTTGAATCCAGCACTTTGGGAG 180
Db 6896 GGTGATTCATAGGTCAGGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAGTTCAGAG 6955

QY 181 TCTACTGAAATATAYAAAA--AGCTAGAGTGGTGGCAGACACCTGTAATCCAGCTAC 236
Db 6956 TCTATTAATAATACAAAAATAGCCAGGATGGTGGCAGGCACTGTAATCCAGCTAC 7015

QY 237 TTAGAGGCTCAGCAGGAGAAATGCTTGAAGCCTAGAGTGAAGGTTGTAGTGAGCCGA 296
Db 7016 TTGGGAGGCTGAGCAGGAGAAATGCTTGAAGCCTGAGGAGTGGAGGTTGCACTGAG 7075

QY 297 GATTGCTCATTTGACAAATGGAGGGAGCCAGCAGCTGGGCAACAGAGAGAAATCTCG 356
Db 7076 GGTGTCACCATG-----GACTCCAGCTGGGCAACAGAGAGAAATCATCG 7121

QY 357 TCTCCAAAAAATAAAAAAAGRAATTAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 416
Db 7122 TCTAAAAAAGAAAAAATAAATAGCAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 7181

QY 417 ACTTGGGAGCAGGG-----GTCCACTTGATGTCGAGACTGCACTGAGCC 462
Db 7182 ACTCAGAGGATGAAGCAAGAGAAATGCTTTGAACCCAGGAGGAGAGATTACAGT 7241

QY 463 ATGATCCTGCCACTGCACTCCGCGCTGGGCAACAGAGTGAAGCCCTGTCTAAAAA 522
Db 7242 GAGATCAACAACACTGCACTCCAGCCTAGGTAA-AGAACAGAGTCCATCTCAAAAA 7300

QY 523 AAAATAAGCAACATATCT 542
Db 7301 TAAATAAATAATATGTTCT 7320
```

```
RESULT 2
US-09-800-960-3
; Sequence 3, Application US/09800960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001158
; CURRENT APPLICATION NUMBER: US/09/800,960
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804
; TYPE: DNA
; ORGANISM: Human
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(62804)
; OTHER INFORMATION: n = A,T,C or G
US-09-800-960-3
```

```
Query Match 14.5%; Score 244.2; DB 4; Length 62804;
Best Local Similarity 73.3%; Pred. No. 3.5e-55;
Matches 407; Conservative 3; Mismatches 101; Indels 44; Gaps 6;

QY 1 GAATCAAGACCAAGCTGGACAACTTGAAGAACCCTGCTCTCTCAAAAAATACAAATT 60
Db 54067 GAGTCAAAACCAAGCTGGCCAAACATGATCAAAACCCGCTCTCTACTAAAAATACAAAA 54126

QY 61 AGCTGGGATTTGGGTGGCGTGCCTCATGCTATAATCCAGCACTTTGGGAGCCTTAGG 120
Db 54127 AGGCC-----GGCGGTAGTGGCTCAGCCCTGTAATCCCAACACTTTGGGAGCGCAAG 54181

QY 121 GGTGGATCACCTGAA--GTCAGGAGTTCAAGACTAGCTGGCCAAACATGGTGAACCCCTAT 179
Db 54182 GGTGGATCACCTGAAGTCAAGAGTTCAGACCCAGCCTGGCCAAACATGGTGAACCTCCAT 54241

QY 180 CTCTACTGAAATATAYAAAAAGCTAGACGTG---GTGGCAGACACCTGTAATCCAGCTAC 236
Db 54242 CTCTACTAAAAATATAAAAAATTAGCCAGGTGTGGGAGGTCCTGTAATCGTAGCTAC 54301

QY 237 TTAGGAGGCTGAGGAGGAGAAATGCTTGAAGCCTAGAGTGAAGGTTGTAGTGAGCCGA 296
Db 54302 TCGGAGGCGGAGGAGGAGGAGTTCGCTTGAACCTGGGAGGTTGCACTGAGCCGA 54361

QY 297 GATTGCTCATTTGACAAATGGAGGGAGCCAGCAGCTGGGCAACAGAGAGAAATCTCG 356
Db 54362 GATCACCCCATTGCA-----CTCAGCCTGGGCAACAGAGCGAAAGTTCT 54407

QY 357 TCTCCAAAAAATAAAAAAAGRAAT-----TAGGCTGGTGGTGGTGGTGGTGGTGG 409
Db 54408 TCTCAAAAAAATAAAAAAATAAATAGCCGGGTGTGGTGGCGGGTCTCTGTAAT 54467

QY 410 CCAGCTACTTGGGAGCGAGGG-----GTCCACTTGATGTCAGACTGCA 455
Db 54468 CCAGCTACTCGGAGACTGAGGCGATGAAATGGCTTGAACCCGAGGAGTGGAGTTGCA 54527

QY 456 GTGAGCATGATCCTGCCACTGCACTCCGCGCTGGGCAACAGAGTGAAGCCCTGTCTAAA 515
Db 54528 GTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGTGAACAGCGAGACTCTGTCTCAA 54587

QY 516 GAAAAAATAATAAA 530
Db 54588 GAAAAAATAATAAAAA 54602
```

```
RESULT 3
US-08-068-945A-1
; Sequence 1, Application US/08068945A
; Patent No. 5616483
; GENERAL INFORMATION:
; APPLICANT: Bjursell, Gunnar
; APPLICANT: Carlsson, Peter
; APPLICANT: Enerback, Sven
; APPLICANT: Hansson, Lennart
; APPLICANT: Lidberg, Ulf
; APPLICANT: Nilsson, Jeanette
; APPLICANT: Tornell, Jan
; TITLE OF INVENTION: New DNA Sequences
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: White & Case
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2787
; COMPUTER READABLE FORM:
```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/068,945A
; FILING DATE: 27-MAY-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9201809-2
; FILING DATE: 11-JUN-1992
; APPLICATION NUMBER: SE 9201826-6
; FILING DATE: 12-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9202088-2
; FILING DATE: 03-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9300902-5
; FILING DATE: 19-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sterner, Richard J.
; REGISTRATION NUMBER: 35,372
; REFERENCE/DOCKET NUMBER: 1103326-052
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)819-8783
; TELEFAX: (212)354-8113
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11531 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: Mammary gland
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
; LOCATION: .4904, 6193..6323, 6501..6608, 6751..6868, 8335
; LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
; FEATURE:
; NAME/KEY: mat\_peptide
; LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
; LOCATION: .4904, 6193..6323, 6501..6608, 6751..6868, 8335
; LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
; OTHER INFORMATION: /EC\_number= 3.1.1.1
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..1640
; FEATURE:
; NAME/KEY: TATA\_signal
; LOCATION: 1611..1617
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1641..1727
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4071..4221
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4307..4429
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4707..4904
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6193..6323
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6501..6608

; FEATURE:
; NAME/KEY: exon
; LOCATION: 6751..6868
; FEATURE:
; NAME/KEY: exon
; LOCATION: 8335..8521
; FEATURE:
; NAME/KEY: exon
; LOCATION: 8719..8922
; FEATURE:
; NAME/KEY: exon
; LOCATION: 10124..10321
; FEATURE:
; NAME/KEY: exon
; LOCATION: 10650..11490
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 11491..11531
; US-08-068-945A-1
;
; Query Match 13.9%; Score 233.6; DB 1; Length 11531;
; Best Local Similarity 71.3%; Pred. No. 1.1e-52;
; Matches 388; Conservative 3; Mismatches 117; Indels 36; Gaps 5;
;
; QY 2 AATTCAAGACCAGCCTGGGACAACTTGGAGAAGCCS---GGTCTCTACAAAAATACAAAA 58
; Db 5164 AGTTCAAGACCAGCCTGAAAATCACTGGGAGAGCCCCCATCTCTACACAAAAATTAANAAT 5223
;
; QY 59 TTAGCTGGGATTGGTCCGGTGGCTCATGCTATATCCAGACACTTTGGGAGCCTGAGG 118
; Db 5224 TAGCTGGGGACTGGCGCGCGGCTCACCTCTGTATATCCAGACAGTGGGAGGCCAAGG 5283
;
; QY 119 TGGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCTTGGCCCAACATGTTGAACCCCTA 178
; Db 5284 TGGGTAGATCACCTGAGGTGAGGAGTTTGAGACAGCCTGACTAAATGGAAGACCTCT 5343
;
; QY 179 TCTCTACTGAAAATAYAAAA--AGCTAGAGCTGGTGGCACACACCTGTATATCCAGCTAC 236
; Db 5344 TCTCTACTAAAAATACAAAAATTAGCCAGGCGTGGTGGCGCTTGCTGTATATCCAGCTAC 5403
;
; QY 237 TTAGGAGGCTGAGCAGGACAATTGCTTGAAGCTAGAGCTGAGGCTTGTAGTGAGCCGA 296
; Db 5404 TCGGGAGGCTGAGGCGAGGAATCGCTTGAACCTCAGGAGCGGAGGTGGCGGTGAGCCGA 5463
;
; QY 297 GATTGTCATTTGCACAATGGAGGGGAGCCACCAGCCTGGGCAACAGAGGAAATCTCCG 356
; Db 5464 GATCATGCCACTGA-----CTCCAGCCTGGAGAACAGAGCTAAACTCTG 5509
;
; QY 357 TCTCCAAAAAATAAAAAAAGRAATTAGGTGG---GTGGTGCCTGTAGTCCCA 413
; Db 5510 TCTCAAAAAAATAAAAAAATAGCCAGCGGTGTATCTCATGCGCTCTGTCTCA 5569
;
; QY 414 GCTACTGGGAGGCGAGGG-----GTCCACTTGATGTCGAGACTGCAGTGA 459
; Db 5570 GCTACTGGGAGGCGAGGTGGGAAGGATCCCTTGAGCCCCAGGGGTTCAAAGTGCAGTGA 5629
;
; QY 460 GCCATGATCTGTCACCTGTCACCTCCGGCTGGGCAACAGAGCTGAGACCCCTGTCTAAAGAAA 519
; Db 5630 GCCGTGTCGTGCCACTGTCACCTCCAGCCTCCAGCCTGGGCGACAGAGTGAGGCCCATCTCAANAAT 5689
;
; QY 520 AAAA 523
; Db 5690 AAGA 5693
;
; RESULT 4
; US-08-442-806-1
; Sequence 1, Application US/08442806
; Patent No. 5716817
; GENERAL INFORMATION:
; APPLICANT: Bjursell, Gunnar
; APPLICANT: Carlsson, Peter
; APPLICANT: Enerback, Sven







```
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
;
US-08-724-394A-20

Query Match 12.8%; Score 215.4; DB 2; Length 246240;
Best Local Similarity 69.9%; Pred. No. 3.3e-47;
Matches 392; Conservative 3; Mismatches 119; Indels 47; Gaps 6;

QY 1 GAATTCAGACCCAGCCTGGACAACTTTGGAGAACCCSGGCTCTTACAAAAAATACAAAAAT 60
|||||
Db 205702 GAATTCAGACCCAGCCTGGACAAACACAGGAGAACCCATCTCTACAAATATACAAAT 205761

QY 61 AGCTGGGATTGGTGGGCTGCTATGCCCTAATATCCAGACACTTTGGGAGCCTGAGGTG 120
|||
Db 205762 AGTGGCC---GAGCGGTGGGCTGACGCGCTGTAATCCAGACACTTTGGGAGCGCGG 205818

QY 121 GTGGGATCACCTGAAGTCAGGAGTTCAAGACTAGCCTGGCCACATGTTGAACCTATC 180
|||
Db 205819 GGCAGATACCTGAGTGAGGATTCACAGCCGCT---CAACATGGAGAACCCCGTC 205875

QY 181 TCTACTGAAAATAYAAAA--AGCTAGACGTGGTGGCACACACCTGTAAATCCCGACTACTT 238
|||||
Db 205876 TCTACTGAAAATACAAAATAGCTGGGCGCTGTGTATGATGCTGCTGTAAATCCCGACTACTC 205935

QY 239 AGGAGCTGAGCAGAGAGATTCCTTGAGCCTGAGGTGAGGTGTAGTGAGCCGAGA 298
|||||
Db 205936 GGGAGGCTGAGGAGAGAGATTCCTTGAGACCTGGGAGGTAGAGTTGCGGTGAGCCGAGA 205995

QY 299 TTGCATCATTTGCACATGGAGGGAGGCCACACAGCCTGGGCAACAGAGAGAAATCTCCGTC 358
|||
```

```
Db 205996 TCCGCGCAATTGCA-----CTCCAGCCTGGGCAACAGGCAAACTCCATC 206041

QY 359 TCACAAAAAATAAAAAAATAAAAAAATAGGTGGGTGGTG-----CCTGTAG 408
|||||
Db 206042 TCACAAAAAATAAAAAAATAAAAAAATAGTCAGGTGTGTGTCACACCTGTAG 206101

QY 409 TCCAGCTACTTGGGAGGCGAGGGGT-----CCACTTGATGTCGAGACTG 453
|||||
Db 206102 TCCAGCTACTTGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCGGGAAGTGTAGGCTA 206161

QY 454 CAGTGAGCCCATGATCCTGCCACTGCACCTCGGCTGGCAACAGAGTGGACCCCTGTCTA 513
|||||
Db 206162 CCATGAGCCCATGATGTCGCACTGTACTCCAGTCTAGGAAAAAATAAACATTAAAT 206221

QY 514 AAGAAAAAATAAAAGCAA 534
|||||
Db 206222 TTAAATCTTAAAAAAGAAA 206242

RESULT 8
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
;
US-08-724-394A-21

Query Match 12.8%; Score 215.4; DB 2; Length 246240;
Best Local Similarity 69.9%; Pred. No. 3.3e-47;
Matches 392; Conservative 3; Mismatches 119; Indels 47; Gaps 6;
```

```
QY 1 GAATTCAGACACCGCTGGACAACTTGGAAAGAACCCSGTCTCTACAAAAATACAAAATT 60
|||||
Db 205702 GAATTCAGACACCGCTGGACAAACACAGGAAAGCCCATCTCTACAAAATATACAAAATT 205761
QY 61 AGCTGGGATTGGGTGGGTGGCTATGCCTATATATCCAGCACCTTTGGGAGCCCTGAGGTG 120
|||||
Db 205762 AGTGGCC---GAGCGGTGGGTGCACGCCCTGTATATCCAGCACCTTTGGGAGCCGAGCG 205818
QY 121 GGTGGATCACCTCAAGTCAAGAGCTTCAAGACTAGCCCTGGGCCAACATGGTGAAACCCCTATC 180
|||||
Db 205819 GGCAGATCACCTGAGTCAGAGGTTCCAGACACGCT---CAACATGGAGAAACCCCGTC 205875
QY 181 TCTACTGAAATAYAAAA---AGTAGACATGGGTGGGCACACACCTGTAAATCCAGCTACTT 238
|||||
Db 205876 TCTACTAAAAATACAAAATTAGCTGGCGGTGGTGATGCATGCCTGTAAATCCAGCTACTC 205935
QY 239 AGGAGCTGAGGAGGAGAAATTCCTTGAAGCCTTAGAGGTGAAGTTGTAGTGAGCCGAGA 298
|||||
Db 205936 GGGAGGCTGAGGAGGAGAAATTCCTTGAACCTGGGAGGTAGAGGTTGCGGTGAGCCGAGA 205995
QY 299 TTGCATCATTCGACAAATGGAGGGAGGAGCCAGCAGCCTGGGCCAACAGAGAAATCTCCGTC 358
|||||
Db 205996 TCCGCCCATTTGCA-----CTCCAGCCTGGGCCAACAGAGCGAAACTCCATC 206041
QY 359 TCCAAAAAATAAAAAAAGRAATTAGCTGGTGGTG-----CCTGTAG 408
|||||
Db 206042 TCAAAAAAATAAAAAAATAAACAATAATAGTCAGGTGTGGTTGTCACACCTGTAG 206101
QY 409 TCCAGCTACTTGGGAGGAGGGGT-----CCACTTGATGTGCAACTG 453
|||||
Db 206102 TCCAGCTACTTGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCGGGAAAGTAGGCTA 206161
QY 454 CAGTGAGCCATGATCCTGCCACTGCACTCCGGCCTGGGCCAACAGAGTGAGACCCCTGTCTA 513
|||||
Db 206162 CCATGAGCCATCATGTGTCACCTGTACTCCAGTCTAGGAAAAAATAAACATTTAAAAATT 206221
QY 514 AAGAAAAAATAAAGCAA 534
|||||
Db 206222 TTAAATCTTAAAAAAGAAA 206242
```

RESULT 9

```
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; NUMBER OF INVENTION: Sequences and Antibodies Thereto
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-Oct-1996
; CLASSIFICATION: 536
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
; US-08-724-394A-22

Query Match 12.8%; Score 215.4; DB 2; Length 246240;
Best Local Similarity 69.9%; Pred. No. 3.3e-47;
Matches 392; Conservative 3; Mismatches 119; Indels 47; Gaps 6;

QY 1 GAATTCAGACACCGCTGGACAACTTGGAAAGAACCCSGTCTCTACAAAAATACAAAATT 60
|||||
Db 205702 GAATTCAGACACCGCTGGACAAACACAGGAAAGCCCATCTCTACAAAATATACAAAATT 205761
QY 61 AGCTGGGATTGGGTGGGTGGCTATGCCTATATATCCAGCACCTTTGGGAGCCCTGAGGTG 120
|||||
Db 205762 AGTGGCC---GAGCGGTGGGTGCACGCCCTGTATATCCAGCACCTTTGGGAGCCGAGCG 205818
QY 121 GGTGGATCACCTCAAGTCAAGAGCTTCAAGACTAGCCCTGGGCCAACATGGTGAAACCCCTATC 180
|||||
Db 205819 GGCAGATCACCTGAGTCAGAGGTTCCAGACACGCT---CAACATGGAGAAACCCCGTC 205875
QY 181 TCTACTGAAATAYAAAA---AGTAGACATGGGTGGGCACACACCTGTAAATCCAGCTACTT 238
|||||
Db 205876 TCTACTAAAAATACAAAATTAGCTGGCGGTGGTGATGCATGCCTGTAAATCCAGCTACTC 205935
QY 239 AGGAGCTGAGGAGGAGAAATTCCTTGAAGCCTTAGAGGTGAAGTTGTAGTGAGCCGAGA 298
|||||
Db 205936 GGGAGGCTGAGGAGGAGAAATTCCTTGAACCTGGGAGGTAGAGGTTGCGGTGAGCCGAGA 205995
QY 299 TTGCATCATTCGACAAATGGAGGGAGGAGCCAGCAGCCTGGGCCAACAGAGAAATCTCCGTC 358
|||||
Db 205996 TCCGCCCATTTGCA-----CTCCAGCCTGGGCCAACAGAGCGAAACTCCATC 206041
QY 359 TCCAAAAAATAAAAAAAGRAATTAGCTGGTGGTG-----CCTGTAG 408
|||||
Db 206042 TCAAAAAAATAAAAAAATAAACAATAATAGTCAGGTGTGGTTGTCACACCTGTAG 206101
QY 409 TCCAGCTACTTGGGAGGAGGGGT-----CCACTTGATGTGCAACTG 453
|||||
Db 206102 TCCAGCTACTTGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCGGGAAAGTAGGCTA 206161
QY 454 CAGTGAGCCATGATCCTGCCACTGCACTCCGGCCTGGGCCAACAGAGTGAGACCCCTGTCTA 513
|||||
Db 206162 CCATGAGCCATCATGTGTCACCTGTACTCCAGTCTAGGAAAAAATAAACATTTAAAAATT 206221
QY 514 AAGAAAAAATAAAGCAA 534
|||||
Db 206222 TTAAATCTTAAAAAAGAAA 206242

RESULT 10
US-09-305-384-5
; Sequence 5, Application US/09305384
; Patent No. 6242218
; GENERAL INFORMATION:
; APPLICANT: Treco, Douglas A.
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Selden, Richard F
```

; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY  
; FILE REFERENCE: 07236/017001  
; CURRENT APPLICATION NUMBER: US/09/305,384  
; CURRENT FILING DATE: 1999-05-05  
; EARLIER APPLICATION NUMBER: US 60/084,649  
; EARLIER FILING DATE: 1998-05-07  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 6235  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-305-384-5

Query Match 12.8%; Score 215.2; DB 4; Length 6235;  
Best Local Similarity 68.5%; Pred. No. 7.2e-48;  
Matches 381; Conservative 2; Mismatches 140; Indels 33; Gaps 5;  
  
QY 2 AATTCAAGACAGCCTGGACAACTTGAAGAACCCSGGTCTCTACAAAAATACAAAAATTA 61  
Db 18 AGTTCAAGACAGCCTGGGAGCAT--AGGAGACTGTCTCTAGAAAAATCAAAAAAT- 74  
  
QY 62 GCTGGGATTGGTGGCTGCTATGCTATATCCAGACACTTTGGGAGCCTGAGGTGG 121  
Db 75 --TATGGCCGGCATGGTGCCTACGCTGTATCCCTGAACCTTTGGGACATCAAGGCAA 132  
  
QY 122 GTGATCACTGGAAGTCAGAGTTCAAGACTAGCCTGGCCAAACATGTTGAACCCCTATCT 181  
Db 133 GTGATCACTTGAGTTCAGAGTTTCAGACTAGCCTGGCCAAACATGTTGAACCCCTATCT 192  
  
QY 182 CTACTGAAAA---TAYAAAAAGCTAGACCTGAGAGTTGAAGTTGTAGTGAGCCGAG 237  
Db 193 CCACATAAAAAATACAAAAATAGCCAGGCATGTTGGCAGCACCTGTATCCCGGTACT 252  
  
QY 238 TAGGAGGCTGAGGAGGAGAAATGCTTGAAGCCTAGAGTTGAAGTTGTAGTGAGCCGAG 297  
Db 253 CAGGAGGCTGAGGAGGAGAAATGCTTGAAGCCTAGAGTTGAAGTTGTAGTGAGCCGAG 312  
  
QY 298 ATTGCATCTTGCACATGAGGAGGAGCCACAGCCTGGGCAACAGAGAGAAATCTCCGT 357  
Db 313 ATCACACCTGCACTCCAGCCTGGGTGACAGAG-----CAAGACTCTATCTCAA 363  
  
QY 358 CTCACAAAAAATAAAAAAAGRATTAGGCTGGGTGGTGGCTGTAGTCCAGCTA 417  
Db 364 AAAAAATAAAAAAATAAAAAATAGCCAGGCATGGTAGTGCACACCTCTAGTCTCAGCTA 423  
  
QY 418 CTTGGGAGCAGGGGGT-----CCACTTGTATGTCGAGACTGCACTGAGCC 462  
Db 424 CTCAGGAGGCTGAGGTGGGAGGATCACTTGAACCTGGGCGAGTCAAGGCTACAGTGAGCC 483  
  
QY 463 ATGATCTGCCACTGCACTCCGGCCTGGGCAACAGAGTGAGACCCCTGTCTAAAGAAAAA 522  
Db 484 AAGATCATGCCACTACACTCCAGCCTGGGCAACAGAGAGAGCCCTGTCTCTAAAAAAT 543  
  
QY 523 AAAATAAGCAACATA 538  
Db 544 AATAATAATAAGAAA 559

RESULT 11  
US-09-305-384-1  
; Sequence 1, Application US/09305384  
; Patent No. 6242218  
; GENERAL INFORMATION:  
; APPLICANT: Treco, Douglas A.  
; APPLICANT: Heartlein, Michael W.  
; APPLICANT: Selden, Richard F.  
; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY  
; FILE REFERENCE: 07236/017001  
; CURRENT APPLICATION NUMBER: US/09/305,384  
; CURRENT FILING DATE: 1999-05-05  
; EARLIER APPLICATION NUMBER: US 60/084,649  
; EARLIER FILING DATE: 1998-05-07

; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 6679  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-305-384-1  
  
Query Match 12.8%; Score 215.2; DB 4; Length 6679;  
Best Local Similarity 68.5%; Pred. No. 7.4e-48;  
Matches 381; Conservative 2; Mismatches 140; Indels 33; Gaps 5;  
  
QY 2 AATTCAAGACAGCCTGGACAACTTGAAGAACCCSGGTCTCTACAAAAATACAAAAATTA 61  
Db 37 AGTTCAAGACAGCCTGGGAGCAT--AGGAGACTGTCTCTACGAAAAATCAAAAAAT- 93  
  
QY 62 GCTGGGATTGGTGGCTGCTATGCTATATCCAGACACTTTGGGAGCCTGAGGTGG 121  
Db 94 --TATGGCCGGCATGGTGCCTACGCTGTATCCCTGAACCTTTGGGACATCAAGGCAA 151  
  
QY 122 GTGATCACTGGAAGTCAGAGTTCAAGACTAGCCTGGCCAAACATGTTGAACCCCTATCT 181  
Db 152 GTGATCACTTGAGTTCAGAGTTTCAGACTAGCCTGGCCAAACATGTTGAACCCCTATCT 211  
  
QY 182 CTACTGAAAA---TAYAAAAAGCTAGACCTGAGAGTTGGTGCACACACCTGTAATCCAGCTACT 237  
Db 212 CCACATAAAAAATACAAAAATAGCCAGGCATGTTGGCAGCACCTGTATCCCGGTACT 271  
  
QY 238 TAGGAGGCTGAGGAGGAGAAATGCTTGAAGCCTAGAGTTGAAGTTGTAGTGAGCCGAG 297  
Db 272 CAGGAGGCTGAGGAGGAGAAATGCTTGAACCCAGGAGCGGAGTTGCACTGAGCTGAG 331  
  
QY 298 ATTGCATCTTGCACATGAGGAGGAGCCACAGCCTGGGCAACAGAGAGAAATCTCCGT 357  
Db 332 ATCACACCTGCACTCCAGCCTGGGTGACAGAG-----CAAGACTCTATCTCAA 382  
  
QY 358 CTCACAAAAAATAAAAAAAGRATTAGGCTGGGTGGTGGCTGTAGTCCAGCTA 417  
Db 383 AAAAAATAAAAAAATAAAAAATAGCCAGGCATGGTAGTGCACACCTCTAGTCTCAGCTA 442  
  
QY 418 CTTGGGAGCAGGGGGT-----CCACTTGTATGTCGAGACTGCACTGAGCC 462  
Db 443 CTCAGGAGGCTGAGGTGGGAGGATCACTTGAACCTGGGCGAGTCAAGGCTACAGTGAGCC 502  
  
QY 463 ATGATCTGCCACTGCACTCCGGCCTGGGCAACAGAGTGAGACCCCTGTCTAAAGAAAAA 522  
Db 503 AAGATCATGCCACTACACTCCAGCCTGGGCAACAGAGAGAGCCCTGTCTCTAAAAAAT 562  
  
QY 523 AAAATAAGCAACATA 538  
Db 563 AATAATAATAAGAAA 578

RESULT 12  
US-09-078-294-7/c  
; Sequence 7, Application US/09078294  
; Patent No. 6265211  
; GENERAL INFORMATION:  
; APPLICANT: Choo, Kong-Hong Andy  
; APPLICANT: Du Sart, Desiree  
; APPLICANT: Cancilla, Michael R.  
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
; FILE REFERENCE: Davies Col  
; CURRENT APPLICATION NUMBER: US/09/078,294  
; CURRENT FILING DATE: 1998-05-13  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 11811  
; TYPE: DNA  
; ORGANISM: BAC-F2 contig 3  
US-09-078-294-7

```
Query Match      12.8%; Score 215; DB 4; Length 11811;
Best Local Similarity 66.1%; Pred. No. 1.1e-47;
Matches 401; Conservative 2; Mismatches 167; Indels 37; Gaps 5;

QY 2 AATTCAAGACGACCTGGACAACTTGAAGAACCSGGTCTCTACAAAAA---TACAAA 57
Db 6599 AGTTCAAGACGACCTGGCCACACAGTAGAGCTCATCTCTACAAAATACATTTAAAG 6540

QY 58 ATTAGCTGGGATTGGGTGGCGTCTCATGCTTATATCCAGACACTTTGGGAGCCTGAG 117
Db 6539 TTAGCTGGGGCCAGGTGTGGTGGCCGACCCCTGTATCCAGCCCTTTGGGAGGTCAAG 6480

QY 118 GTGGGTGATCACTGAAGTCAGAGTTCAAGACTAGCTAGCTGGCCACACATGGTGAACCC 177
Db 6479 GTGGGTGATCACTGAAGTCAGAGTTCAAGACTAGCTAGCTGGCCACACATGGTGAACCC 6420

QY 178 ATCTCTACTGAAATAYAAAA---AGCTAGAGCTGGTGGCCACACACACTCTAATCCAGC 233
Db 6419 ATCTCTACTGAAATAYAAAAATTAGCCAGCGGTGGTGGCGCGCTGTATATCCAGC 6360

QY 234 TACTTAGAGGCTGAGGACGAGAAATTTGCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGC 293
Db 6359 TACTCAGAGGCTGAGGACGAGAAATCACTTGAACCCAGAGAGAGGTGTAGTGAGC 6300

QY 294 CGAGATTGCATATGACAAATGGAGGGAGCCACAGCTGGGCAACAGAGGAAATCT 353
Db 6299 CAAGATCATGCCATTGCA-----CTCCAGCCTGGGCAACAGAGCAAACT 6254

QY 354 CCGTCTCCAAAAAATAAAAAAAGRAATTAGGCTGGGTGGTGGCTGTAGTCCCA 413
Db 6253 CTCAAAAAATAAAAAAATAAAATTTCCCGGGGTATGGGGGCAATTCCTGTAGAGGTA 6194

QY 414 GCTACTTGGGAGGCGAGGGGTCCA-----CTTGATGTCGAGACTGCAGTGA 459
Db 6193 GCTACATGGAAGCTGAGGACGAGGAGTTGCTTGGGTCCAGGATTTGAGGTTACAGTGA 6134

QY 460 GCATG-ATCCTGGCACTGCACCTCCGGCCTGGGCAACAGAGTGAGACCCCTGTCTAAAGAA 518
Db 6133 GATACGAAATGGTCCCACTGCATACAGCTTGGGAGACAGCTTGTCTCTAAAAAAA 6074

QY 519 AAAAAAATAAAGCAACATATCTGAACAAAGGATCCTCCATACGPTTCCACACAGATT 578
Db 6073 AATAGTAAAAAACAACAACAAACCCCTGGCTGTGCATGGTGGCTCACCCCTGTAA 6014

QY 579 CTAATCA 585
Db 6013 CCAACA 6007

RESULT 13
US-07-914-281-5/c
; Sequence 5, Application US/07914281
; Patent No. 5324663
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/914,281
; FILING DATE: 19920720
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8174 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; US-07-914-281-5

Query Match      12.8%; Score 214.6; DB 1; Length 8174;
Best Local Similarity 68.0%; Pred. No. 1.2e-47;
Matches 379; Conservative 1; Mismatches 155; Indels 22; Gaps 5;

QY 1 GAATTCAGACGACCTGGACAACTTTGGAAGAACCSGGTCTCTACAAAAAATACAAAATT 60
Db 4362 GAGTTCAAGATCAGCCCTGGGCAACAGAGCA-CTCTTACAAAAAATTTTAAATTAGCTT 4304

QY 61 AGCTGGGATTGGGTGGCGTGCATGCCTATATATCCAGCACTTTGGGAGCCCTGAGGTG 120
Db 4303 GGCATGGGCCAGCGCGGTGGCTCACACCTGTAAATCCACCACTTTGGGAGGCCAAGGTG 4244

QY 121 GGTGGATCAGCTCAAGTCAGAGTTCAAGACTAGCCTGGCCACACATGGTGAACCCCTATC 180
Db 4243 GGTGGATCAGCTCAAGTTGGGAGTTTCGAGACCAGCCTTGACCAACGTTGGAGAAACCCCTGC 4184

QY 181 TCTACTGAAAAATAYAAAA--AGCTAGACGTGGTGGCACACACCTGTAAATCCCACTACTT 238
Db 4183 TCTACTTAAATACAAATTTAGCCGGGCATGGTGGCGCATGCCTGTAAATCCAGCCACTC 4124

QY 239 AGAGGCTGAGGACGAGAAATTTGTTGAAGCCCTAGAGGTGAAGGTGTAGTGAGCCGAGA 298
Db 4123 GGGAGGCTGAGGACGAGAAATCGCTGAACCCGGGGGGGAGTTGCGGTGAGCTGAGA 4064

QY 299 TTGCATCATTCACAATGGAGGGGAGCCACAGCCTGGGCAACAGAGGAAATCTCCGTC 358
Db 4063 TCATGCCATTACA-----CTCCAGCCTGGGCAACAGAGTGAAACTCCGTC 4018

QY 359 TCCAAAAAATAAAAAAAGRAATT-AGGCTGGGTGGTGGCTGTAGTCCAGCTA 417
Db 4017 TCCAAAAAATAAATAAATTTAGCTTGGCATGGTGGCACATGTCTGTGCTCAGCTA 3958

QY 418 CTTGGGAGGCGGGGTCCACTTGTATGTCGAGACTCGAGTGAGCCATGATCTCCGTC--- 473
Db 3957 CACCGGATGTAAAGCGGGGAGGATCCCGGAGCTCAACAATGAGCCGATAGCACCGCTG 3898

QY 474 ACTGCACTCCGGCCTGGGCAACAGAGTGAGACCCCTGCTCTAAAGAAAAAATAAAGCA 533
Db 3897 ACTGCACTCCAGCTTGGCGACAGAGAGGAGCCCTGCTCTTAAAAAATAAAAAA 3838

QY 534 ACATATCCTCGAACAAAG 550
Db 3837 AAAGAAAGTGGTCCAGG 3821

RESULT 14
US-08-393-246-5/c
; Sequence 5, Application US/08393246
; Patent No. 5595900
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
```



```
QY 121 GGTGATCACCTGAAGTCAAGGATTCAAGACTAGCCTGGCCACACATGGTGAACCCCTATC 180
Db 4243 GGTGGATCACCTGAGGTTGGGAGTTTCGAGACCAGCCTGACCAACGTTGGAGAAACCCCTGTC 4184
QY 181 TCTACTGAAATAYAAAA--AGCTAGACGTGGTGGCACACACCTGTAAATCCCGAGCTACTT 238
Db 4183 TCTACTTAAATACAAAATTAGCCGGGCATGGTGGCGCATGCCCTGTAAATCCCGAGCCACTC 4124
QY 239 AGGAGCTGAGCGAGGAGAAATTGCTTTGAAGCCTAGAGGTGAAGGTTGTAGTGAGCGGAGA 298
Db 4123 GGGAGCTGAGCGAGGAGAAATCGCCTGAACCCGGGGGGGGAGTTGCCGTGAGCTGAGA 4064
QY 299 TTGCATCATTTGCACAAATGGAGGGGAGCCACCGAGCCTGGGCAACAAGAGGAAATCTCCGTC 358
Db 4063 TCATGCCATTACA-----CTCCAGCCTGGGCAACAAGAGTGAAACTCCGTC 4018
QY 359 TCCAAAAAATAAAAAAAGRATT-AGGCTGGGTGGTGGCTGTAGTCCAGCTA 417
Db 4017 TCCAAAAAATAAATAAATTTAGCTTGGCATGGTGGCACAATGTCTGTGGTCTCAGCTA 3958
QY 418 CTTGGGAGCGAGGGGTCCACTTGTATGTCGAGACTGCAGTGAGCCATGATCCTGCC- --- 473
Db 3957 CACCGGATGCTAAGGCGGGAGGATCCCGGAGCTCAAAATGAGCCGCGATAGCACCCTG 3898
QY 474 ACTGCACCTCCGGCTGGGCAACAGAGTGAGACCCCTGTCTAAAGAAAAAATAAAGCA 533
Db 3897 ACTGCACCTCCAGCTTCCGCGACAGAGAGGGACCCCTCTCTTAAAAAATAAAAAA 3838
QY 534 ACATATCCTGACAAAG 550
Db 3837 AAAGAAAGTGGTCCAGG 3821
```





Result No.	Score	Query		Length	DB	ID	Description
		Match					
C 1	249	14.8	57130	10	US-09-835-081-3		Sequence 3, Appli
C 2	244.2	14.5	62804	12	US-10-096-960-3		Sequence 3, Appli
C 3	240.2	14.3	15849	10	US-09-880-107-2362		Sequence 2362, Ap
C 4	240.2	14.3	23470	10	US-09-764-847-1157		Sequence 1157, Ap
C 5	238.8	14.2	32190	10	US-09-764-869-2209		Sequence 2209, Ap
C 6	237.8	14.2	65608	9	US-09-954-531-180		Sequence 180, App
C 7	237.8	14.2	65608	10	US-09-962-436-292		Sequence 292, App
C 8	237.8	14.2	65608	10	US-09-962-832-119		Sequence 119, App
C 9	237.6	14.1	4963	10	US-09-764-877-2903		Sequence 2903, Ap
C 10	235.6	14.0	84539	10	US-09-962-436-36		Sequence 36, Appli
C 11	233.8	13.9	465237	9	US-09-933-2678-A-1		Sequence 1, Appli
C 12	232.8	13.9	15297	9	US-10-003-295-3		Sequence 3, Appli
C 13	232.4	13.8	110096	10	US-09-880-107-1542		Sequence 1542, Ap
C 14	231.6	13.8	5386	10	US-09-764-864-1661		Sequence 1661, Ap
C 15	231.4	13.8	99014	10	US-09-880-107-3428		Sequence 3428, Ap
C 16	230.8	13.7	3966	10	US-09-764-855-227		Sequence 227, App
C 17	230.2	13.7	16086	10	US-09-764-877-2385		Sequence 2385, Ap
C 18	229.8	13.7	9970	10	US-09-764-877-2277		Sequence 2277, Ap
C 19	229.8	13.7	10195	10	US-09-764-864-1600		Sequence 1600, Ap





## RESULT 6

US-09-954-531-180  
; Sequence 180, Application US/09954531  
; Patent No. US20020165180A1

## GENERAL INFORMATION:

; APPLICANT: Weaver, Zoe  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand

; FILE REFERENCE: 689290-77

; CURRENT APPLICATION NUMBER: US/09/954,531

; CURRENT FILING DATE: 2002-05-02

; PRIOR APPLICATION NUMBER: US/60/233,133

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,009

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,034

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,509

; PRIOR FILING DATE: 2000-09-22

; PRIOR APPLICATION NUMBER: US/60/234,567

; PRIOR FILING DATE: 2000-09-22

; NUMBER OF SEQ ID NOS: 1392

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 180

; LENGTH: 65608

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(65608)

; OTHER INFORMATION: n=a,t,g or c

US-09-954-531-180

Query Match 14.2%; Score 237.8; DB 9; Length 65608;

Best Local Similarity 69.7%; Pred. No. 2.3e-43;

Matches 388; Conservative 3; Mismatches 135; Indels 31; Gaps 4;

QY 3 ATTCACAGCCAGCGTGGACAACTTTGGAAGAACCGSGTCTCTACAAAAATACAAAATTAG 62  
DB 39025 ATTCCACTCCAGCCTGGGAAGAGCGAGACTTTGTCTCCAAAAAATAAAAAAATTT 39084  
QY 63 CTGGGATTGGGTGGGTGCGTCATGCCCTAATAATCCAGCACTTTGGAGCGCTGAGGTGG 122  
DB 39085 GGCAGGCCAGGCACAGTGGCTCACACCTGTAATCCAGCCCTCTGGGAGCCGAGCAG 39144  
QY 123 TGGATCACCTGAAGTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAACCCATCTC 182  
DB 39145 AGGATCTCTGAGGTCAGGAGTTTGAGAACAGCCTGACATAGTGAACCCCATCTC 39204  
QY 183 TACTGAAAATAYAAAA--AGCTAGACGTGGTGGCACACACCTGTAATCCAGCTACTTAG 240  
DB 39205 TACTAACATAAATAATAGCCAGGTGTGATGGCACATGCCCTGAATCCAGCTACTTG 39264  
QY 241 GAGGCTGAGCGAGGAGAAATTTGTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCCGAGAT 300  
DB 39265 GGGGTTGAGCGAGGAGAAATTTGTTGAACCCAGGAGCAGAGTTGCACTGAGCCGAGATC 39324  
QY 301 GCATCATTCACAATGAGGGGAGCCAGCAGCTGGGCCAACAGAGAAATCTCCGCTCTC 360  
DB 39325 GCACCATTTGCA-----CCCGAGCCTGGGCACACAGCGGAACCTCCATCT- 39369  
QY 361 CAAAAAATAAAAAAAGRAATTAGGCTGGGTGGTGTAGTCCCGAGCTACTT 420  
DB 39370 CAAAAAATAAAAAAATAGTTGGGCATGGTGGCATGCACCTATAGTCCCGAGCTACTC 39429  
QY 421 GCGAGCGAGGG-----GTCCACTTGATGTCGAGACTGCGAGTGAGCCATGA 466  
DB 39430 AGGAGGCTGAGGTGGGAGGATCTTTGAGCCCAAGAGATCAAGGCTGCGAGTGGCCATGT 39489  
QY 467 TCCTGCCACTGCACTCCGCGCTGGGCAACAGAGTGAGCCCTGTCTAAAGAAAAA 526  
DB 39490 TTGCACCACTGCACTCCAGCCTGGGCAACAAAGACTCTGTCTCAAAAAA 39549

Db 39490 TTGCACCACTGCACTCCAGCCTGGGCAACAAAGACTCTGTCTCAAAAAA 39549

QY 527 TAAGCAACATATCCTG 543

Db 39550 AAAAAAAAAAAGGCAG 39566

## RESULT 7

US-09-962-436-292

; Sequence 292, Application US/09962436

; Patent No. US20020081301A1

; GENERAL INFORMATION:

; APPLICANT: Soppet, Daniel

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign

; TITLE OF INVENTION: Sets

; FILE REFERENCE: 689290-75

; CURRENT APPLICATION NUMBER: US/09/962,436

; CURRENT FILING DATE: 2001-09-25

; PRIOR APPLICATION NUMBER: US/60/235,082

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/234,924

; PRIOR FILING DATE: 2000-09-25

; NUMBER OF SEQ ID NOS: 568

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 292

; LENGTH: 65608

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; OTHER INFORMATION: n=a,t,g or c

US-09-962-436-292

Query Match

Best Local Similarity 69.7%; Score 237.8; DB 10; Length 65608;

Matches 388; Conservative 3; Mismatches 135; Indels 31; Gaps 4;

QY 3 ATTCACAGCCAGCGTGGACAACTTTGGAAGAACCGSGTCTCTACAAAAATACAAAATTAG 62  
DB 39025 ATTCCACTCCAGCCTGGGAAGAGCGAGACTTTGTCTCCAAAAAATAAAAAAATTT 39084  
QY 63 CTGGGATTGGGTGGGTGCGTCATGCCCTAATAATCCAGCACTTTGGAGCGCTGAGGTGG 122  
DB 39085 GGCAGGCCAGGCACAGTGGCTCACACCTGTAATCCAGCCCTCTGGGAGCCGAGCAG 39144  
QY 123 TGGATCACCTGAAGTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAACCCATCTC 182  
DB 39145 AGGATCTCTGAGGTCAGGAGTTTGAGAACAGCCTGACATAGTGAACCCCATCTC 39204  
QY 183 TACTGAAAATAYAAAA--AGCTAGACGTGGTGGCACACACCTGTAATCCAGCTACTTAG 240  
DB 39205 TACTAACATAAATAATAGCCAGGTGTGATGGCACATGCCCTGAATCCAGCTACTTG 39264  
QY 241 GAGGCTGAGCGAGGAGAAATTTGTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCCGAGAT 300  
DB 39265 GGGGTTGAGCGAGGAGAAATTTGTTGAACCCAGGAGCAGAGTTGCACTGAGCCGAGATC 39324  
QY 301 GCATCATTCACAATGAGGGGAGCCAGCAGCTGGGCCAACAGAGAAATCTCCGCTCTC 360  
DB 39325 GCACCATTTGCA-----CCCGAGCCTGGGCACACAGCGGAACCTCCATCT- 39369  
QY 361 CAAAAAATAAAAAAAGRAATTAGGCTGGGTGGTGTAGTCCCGAGCTACTT 420  
DB 39370 CAAAAAATAAAAAAATAGTTGGGCATGGTGGCATGCACCTATAGTCCCGAGCTACTC 39429  
QY 421 GCGAGCGAGGG-----GTCCACTTGATGTCGAGACTGCGAGTGAGCCATGA 466  
DB 39430 AGGAGGCTGAGGTGGGAGGATCTTTGAGCCCAAGAGATCAAGGCTGCGAGTGGCCATGT 39489  
QY 467 TCCTGCCACTGCACTCCGCGCTGGGCAACAGAGTGAGCCCTGTCTAAAGAAAAA 526  
DB 39490 TTGCACCACTGCACTCCAGCCTGGGCAACAAAGACTCTGTCTCAAAAAA 39549

```
QY 527 TAAAGCAACATATCCTG 543
Db 39550 AAAAAAAAAAAGGCAG 39566

RESULT 8
US-09-962-832-119
; Sequence 119, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sels
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 119
; LENGTH: 65608
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: n,a,t,g or c
US-09-962-832-119

Query Match 14.2%; Score 237.8; DB 10; Length 65608;
Best Local Similarity 69.7%; Pred. No. 2.3e-43;
Matches 388; Conservative 3; Mismatches 135; Indels 31; Gaps 4;

QY 3 ATTCAAGACCGCTGGACAACTTGGAGAACCGSGGTCTCTACAAAAATACAAAAATTAG 62
Db 39025 ATTCCATCCAGCGCTGGGAAGAGCGAGACTTTGTCTCCAAAAAATAAAAAAATAATT 39084

QY 63 CTGGGATTGGGTGGGTGGCTCATGCCCTATATCCAGCACTTTGGGAGCGCTGAGGTGGG 122
Db 39085 GGCAGGCCAGGCACAGTGGCTCACACTGTATCCAGCCCTCTGGGAGCGCGAGCAGG 39144

QY 123 TGGATCACTGAAGTCAAGAGTTCAAGACTAGCGTGGCCCAACATGGTGAACCCATATCTC 182
Db 39145 AGGATCTCTCAGTCAAGAGTTTGAACACAGCGCTGACTGACATAGTGAACCCCATCTC 39204

QY 183 TACTGAAATAYAAAA--AGCTAGACGTGGTGGCACACACCTGTAAATCCAGCTACTTATG 240
Db 39205 TACTAACAATACAAAAATTAGCCAGGTGTGATGGCAGCATGCCCTGAAATCCAGCTACTTGG 39264

QY 241 GAGGCTGAGCGAGGAGAAATTGCTTGAAGCCCTAGAGGTGAAGTTGTAGTGAGCGCGAGATT 300
Db 39265 GGGGTTGAGCGAGGAGAAATTGCTTGAACCCAGGAGCAGAGTTGAGTGAAGCGCGAGATC 39324

QY 301 GCATCATTTGCACATGAGGAGGAGCCACACGCTGGGCAACAAGAGGAATCTCCGTCCTC 360
Db 39325 GCACCAATTGCA-----CCCCAGCCTGGGCAACAAGAGGGAATCCCATCT- 39369

QY 361 CAAAAAATAAAAAAATAAAAAAAGRATTAGCTGGGTGGTGGCTGTAGTCCAGCTACTTT 420
Db 39370 CAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 39429

QY 421 GGGAGGAGGAGG-----GTCCACTTTGATGTGCGAGACTGCACTGAGCCATGA 466
Db 39430 AGGAGGCTGAGGTGGGAGGATCTTTGAGCCCAAGAGATCAAGGCTGCATGAGCCATGT 39489

QY 467 TCCTGCCACTGCACCTCGGCGCTGGGCAACAGAGTGAAGCCCTGTCTATAAGAAAAAATAAAA 526
Db 39490 TTGCACCACCTGCACCTCCAGCTGGGCAACAACAAAGACTCTGTCTCAAAAAAATAAAAAA 39549

QY 527 TAAAGCAACATATCCTG 543
Db 39550 AAAAAAAAAAAGGCAG 39566
```

```
Db 39550 AAAAAAAAAAAGGCAG 39566

RESULT 9
US-09-764-877-2903
; Sequence 2903, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2903
; LENGTH: 4963
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2903

Query Match 14.1%; Score 237.6; DB 10; Length 4963;
Best Local Similarity 71.6%; Pred. No. 1.2e-43;
Matches 398; Conservative 3; Mismatches 107; Indels 48; Gaps 5;

QY 1 GAATTCAAGACCAACCGCTGGACAACTTGGAAAGAACCCSGTCTCTACAAAAAATACAAAAATT 60
Db 828 GAGTTCAAGACCAACCGCTGGGCAACATGGTGAACCCCTGTCCCTACTAAATAACACACAC 887

QY 61 AGCTGGGATTGGGTGGGTGGCTCATGCCCTATATATCCAGCACTTTTGGGAGCGCTGAGGTG 120
Db 888 AGGCC-----GGCGCGGTGGCTTACGCCCTGTAAGCCAGCACTTTGGGAGCGCGAGGAG 942

QY 121 GGTGGATCACTCAAGTCAGGAGTTCAAGACTAGCCCTGGGCCAACATGTTGTAACCCCTATC 180
Db 943 GGCAGATCACTCAGGTGGGAGTTCAAGACCAAGCTGACCAACATGAGAGAAACCCCGTC 1002

QY 181 TCTACTCAAAATAYAAAA--AGCTAGACGTGGTGGCACACACACCTGTAAATCCAGCTACTT 238
Db 1003 TCTACTAAATAACAAAAATTAGCCGCGGTGGTGGCCATGCCCTGTAAATCCAGCTACTC 1062

QY 239 AGAGGCTGAGGAGGAGAAATTGCTTGAAGCCCTAGAGGTGAAGGTGTAGTGAGCCGAGA 298
Db 1063 GGGAGGCTGAGGAGGAGAAATCGCTTGAACCCGGAGCGGAGGTGTGTGAGCTGAAA 1122

QY 299 TTGCATCATTTGCACAAATGGAGGGAGCCACACGCTGGGCCAACAGAGAAATCTCCGTC 358
Db 1123 TCGTGCCATTGCA-----CTCCAGCCTGGGCCAACAGAGCGAAACTCCATC 1168

QY 359 TCCAAAAAATAAAAAA-----AAAAAAGRATTAGGCTGGGTGGTGGCTG 405
Db 1169 TCAAAAAAATAAAAAAATAAGCTACACACACACAAAAATACAGGCATGGGGTGCATGCGCTG 1228

QY 406 TAGTCCAGCTACTTGGGAGCGAGGG-----GTCCACTTGTATGTCGAGAC 451
Db 1229 TAATCCAGCTACTTGGGAGCGCTGAGGCATGAGCATCGCTTGAACCTGGGAGCGGAGGT 1288

QY 452 TGCAGTGAGCCATGCTCCACTGCACTCCGCGCTGGGCCAACAGAGTGAAGCCCTGTC 511
Db 1289 TGCAGTGAGCCATGCTCCACTGCACTTACCGCTGACCTAGGTGACGGAGTGAGATTGTGTC 1348

QY 512 TAAAGAAAAAATAAAT 527
Db 1349 TCCAAAAAATAAAAAAT 1364

RESULT 10
US-09-962-436-36/c
; Sequence 36, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
```

```

; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
;
; TITLE OF INVENTION: Sets
;
; FILE REFERENCE: 689290-75
;
; CURRENT APPLICATION NUMBER: US/09/962,436
;
; CURRENT FILING DATE: 2001-09-25
;
; PRIOR APPLICATION NUMBER: US/60/235,082
;
; PRIOR FILING DATE: 2000-09-25
;
; PRIOR APPLICATION NUMBER: US/60/234,924
;
; PRIOR FILING DATE: 2000-09-25
;
; NUMBER OF SEQ ID NOS: 568
;
; SOFTWARE: PatentIn version 3.0
;
; SEQ ID NO 36
;
; LENGTH: 84539
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; US-09-962,436-36

```

```

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 15297
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-003-295-3

Query Match      13.9%; Score 232.8; DB 9; Length 15297;
Best Local Similarity 69.0%; Pred. No. 1.9e-42;
Matches 381; Conservative 3; Mismatches 145; Indels 23; Gaps 4;

Qy 1 GAATTCAGACCGCTGGGACAACTTGGAGAAACGSGTCTCTACAAAAATACAAAAAT 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11765 GAGTTCAAGATCAGCTTGGGACACACAGTAGAACTCCACTGTGTACAAAAATACAAAAAT 11824

Qy 61 AGCTGGGATTCGGGTGGCTCATGCTCTAATATCCAGCACCTTTGGAGCCTGAGGTG 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11825 A-----GACTGGGACGGTGGCTCACACTGTATCCAGCACCTTTGGAGGCGGAGGCA 11879

Qy 121 GGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCCTTGGCCCAACATGGTGAACCCCTATC 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11880 GGTGGATCACCTGGTTCAGGAGTTTGAGACCGCAGCCAGCAACATGGTGAACCCCATC 11939

Qy 181 TCTACTGAAATATAAAA--AGCTAGACGTGGTGGGCACACACCTGTATCCAGCTACT 237
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 11940 TCTACTAAAAATACAAAAATTTAGCCAGGCATGGTGGCACGTGCCTGTATCCAGCTACT 11999

Qy 238 TAGGAGGCTGAGGAGGAGAAATTCCTTCAAGCCTAGAGGTCAAGGTCTAGTGAGCCGAG 297
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12000 TGGAGGCTGAGGTGGGAGAAATTCCTTGAACCCAGGAGCGGAGGCTGCAGTGAGCCGAG 12059

Qy 298 ATTGCATCATTTGCACAATAGGAGGGAGCCACAGCCTGGGCAACAGAGGAAATCTCCGT 357
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12060 ATTGTGCACCTGCACCTCCAGCCTGG--GGCAAGAGTGAATCCATCTCAAAAAAAC 12118

Qy 358 CTCCAAAAAAATAAAAAAAGRATTAGCTGGGTGGTCTAGTCCAGCTA 417
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12119 AAAAAAATAAAAAATACAAAAATTTAGCTGGGTGTGGTGCATCGCCTGTAGTCCCTGCTA 12178

Qy 418 CTGGGAGGCGAGG-----GGTCCACTTGATTCGAGACTCGAGCTGAGTGAGCA 463
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12179 CTCGGGAGGCTGAGGTGGGAGGACTCTGGAGCCCGGAGGTGGAGGTTCAGTGAGCTG 12238

Qy 464 TGATCTCTGCCACTGCACCTCCGCGCTGGGCAACAGAGTGCAGCCCTGTCTAAGAAAAAAA 523
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 12239 AGATCATGCCACTGCACCCCAACCTGGGTGCACAGAGAGAGAGACCTTGCCTCGAAA 12298

Qy 524 AATAAAGCAAC 535
   ||||| ||||| |||||
Db 12299 AAAAAAACC 12310

RESULT 13
US-09-880-107-1542/c
; Sequence 1542, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1542

```

Db	976	TGGGGCTGGGCACAGTGGCTCAGCTGTGTAATCCACGACCTTTGGGAAGCGGAATAGGT	1035
Qy	124	GGATCACCTGAAGTCAAGGATTCAGAGCTAGGCTGCCAACATGGTGAAGCCCTATCTCT	183
Db	1036	GGATCCTAGAGGTCTATGAGTTCAACACACACCCCTGGCCAACATGCAGAAATCCATCTCT	1095
Qy	184	ACTGAAATAYAAAA-----AGCTAGACGTGGTGGGCACACCTGTGTAATCCAGCATCTT	238
Db	1096	ATTAAAAATACAAAAACAGTAGCCGGGTGTGGTGGCGCATGCCTGTAATCCACGACTATG	1155
Qy	239	AGGAGGCTGAGCCAGGAGAAATTCCTTGAAGCCTCAGAGTGAAGGTGTGAGTGAGCCGAGA	298
Db	1156	GGGAGGCTGAGGCACGACGAGATTCGCTTGAACCTGGAGGTGGAGGTGTGGTGAGCCCTAGA	1215
Qy	299	TTGCATCATTTGCACATGAGGAGGAGCCACCAGCCTGGGCCAACAGAGGAAATCTCCGTC	358
Db	1216	TTGTACCACATGCA-----CTTCAGCCTGGCAACGAGAGCGAAATCCGTC	1261
Qy	359	TCCAAAAAATAAAAAAAAAAAGRATTAGGCTGGGTGGTGCCTGTAGTCCAGCTAC	418
Db	1262	TCATAAAAAAAAAAATAACAAAAACAGGTGGTGGTGTCACACTGTGTAATTCACAGTTC	1321
Qy	419	TTGGGAGGCAGGG-----GTCCACTTGATGTCGAGACTGCAGTGAGCCAT	464
Db	1322	TTGGGAGGCTGAGGCATGAGAACTCACTTGAACCTGGGAGGCAGAGGTTCGAGTGACCTGA	1381
Qy	465	GATCTGCCACTGTCACCTCCGGCTGGCAACAGAGTGAGACCCCTGTCTAAGAAAAAATA	524
Db	1382	GAGAGTGCCNCTGCATCCAGCTCCAGCCTGGCAACAAAGTGAGACTGTCTCTCAAAACAAAA	1441
Qy	525	AATAAGCAAA	534
Db	1442	ACAAAAAGAAA	1451

## RESULT 15

US-09-880-107-3428/c

```

: Sequence 3428, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3428
: LENGTH: 99014
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: Genbank Accession No. US20020142981A1 U7
: US-09-880-107-3428

```

Query Match 13.8%; Score 231.4; DB 10; Length 99014;  
Best Local Similarity 71.1%; Pred. NO. 6.8e-42;  
Matches 389; Conservative 2; Mismatches 119; Indels 38; Gaps 5;

[illegible]



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 08:36:11 ; Search time 4501 Seconds  
(without alignments)  
10862.634 Million cell updates/sec

Title: US-09-942-310-2

Perfect score: 1680

Sequence: 1 gaattcaagaccagctgga.....catcttcctgctcctgggtgg 1680

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 20

*BY MOVE*

Total number of hits satisfying chosen parameters: 233324

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

GenEmbl:

1: gb\_ba:

2: gb\_hgt:

3: gb\_in:

4: gb\_om:

5: gb\_ov:

6: gb\_pat:

7: gb\_ph:

8: gb\_pl:

9: gb\_pr:

10: gb\_ro:

11: gb\_sts:

12: gb\_sy:

13: gb\_un:

14: gb\_vi:

15: em\_ba:

16: em\_fun:

17: em\_hum:

18: em\_in:

19: em\_mu:

20: em\_om:

21: em\_or:

22: em\_ov:

23: em\_pat:

24: em\_ph:

25: em\_pl:

26: em\_ro:

27: em\_sts:

28: em\_un:

29: em\_vi:

30: em\_hgt\_hum:

31: em\_hgt\_inv:

32: em\_hgt\_other:

33: em\_hgt\_mus:

34: em\_hgt\_pln:

35: em\_hgt\_rod:

36: em\_hgt\_mam:

37: em\_hgt\_vrt:

38: em\_sy:

39: em\_hgtgo\_hum:

40: em\_hgtgo\_mus:

41: em\_hgtgo\_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1680	100.0	1680	6	AX394457	AX394457 Sequence
2	1516	90.2	1669	6	AX207224	AX207224 Sequence
3	1338	79.6	9432	6	AX394456	AX394456 Sequence
4	1338	79.6	9432	9	HUMCYP2D6	M33388 Human cytoC
5	1158	68.9	13677	9	HSCYP2D7B	X58468 Human CYP2D
6	721	42.9	5503	9	HUMCYP2DGA	M33189 Human debri
7	287	17.1	13278	9	HSCYP2D7A	X58467 Human CYP2D
8	269	16.0	17060	9	HUMCYP8P	M33387 Human debri
9	269	16.0	114846	9	HS257120	AL021878 Human DNA
C	138	8.2	1593	6	AX195173	AX195173 Sequence
10	89	5.3	1563	9	HSP450IID	X16867 Human mRNA
11	61	3.6	1494	6	E10647	E10647 Human CDNA
12	61	3.6	1494	6	E10867	E10867 cDNA encod1
13	61	3.6	1494	6	E10868	E10868 CDNA encod1
14	61	3.6	1494	6	E10869	E10869 CDNA encod1
15	61	3.6	1494	6	E10870	E10870 CDNA encod1
16	61	3.6	1566	6	AR084365	E10870 CDNA encod1
17	61	3.6	1566	6	AR084374	AR084365 Sequence
18	61	3.6	1567	6	AR084374	AR084374 Sequence
19	61	3.6	1567	9	HSD81	X08006 Homo sapien
20	61	3.6	1567	9	HUMCYPDB1	M20403 Human cytoC
21	61	3.6	1568	6	A20907	A20907 debri
22	61	3.6	1568	6	AR084372	A20907 debri
23	61	3.6	1571	6	AR084373	AR084372 Sequence
24	61	3.6	172611	9	AC019129	AC019129 Homo sapi
C	60	3.6	155401	9	AC007551	AC007551 Homo sapi
25	60	3.5	200807	9	AC007551	AC007551 Homo sapi
26	59	3.5	190349	9	AC009475	AC009475 Homo sapi
C	58	3.5	190349	9	AC009475	AC009475 Homo sapi
27	58	3.3	107415	9	AL513365	AL513365 Homo sapi
28	56	3.3	173519	2	AL732423	AL732423 Homo sapi
C	55	3.3	163111	9	AC080014	AC080014 Homo sapi
30	55	3.3	176232	9	AC011825	AC011825 Homo sapi
C	55	3.3	176232	9	AC011825	AC011825 Homo sapi
31	55	3.3	178834	9	CNS07EFF	AL512791 Human chr
32	55	3.3	185402	9	AC006536	AC006536 Homo sapi
C	55	3.3	214984	9	AC015563	AC015563 Homo sapi
33	55	3.3	228097	9	CNS00YVH	AL096869 Human chr
C	55	3.3	228097	9	CNS00YVH	AL096869 Human chr
35	55	3.2	265308	9	AF228730	AF228730 Homo sapi
36	54	3.2	94850	9	AC025275	AC025275 Homo sapi
C	53	3.2	114411	9	HUAC002551	AC025275 Homo sapi
37	53	3.2	114411	9	HUAC002551	AC025275 Homo sapi
38	53	3.2	149628	2	AC034128	AC02551 Human chr
39	53	3.2	168373	9	AL391683	AC034128 Homo sapi
40	53	3.2	178378	9	AC025821	AL391683 Human DNA
41	53	3.2	181438	2	AC023047	AC025821 Homo sapi
42	53	3.2	181438	2	AC023047	AC023047 Homo sapi
43	53	3.2	182908	2	AC127070	AC127070 Homo sapi
C	53	3.2	303014	2	AC117448	AC127070 Homo sapi
44	53	3.1	100000	9	AP000086	AC117448 Homo sapi
C	52	3.1	123943	9	AC006208	AP000086 Homo sapi
46	52	3.1	130069	9	AC044839	AC006208 Homo sapi
47	52	3.1	142323	9	AC010988	AC006208 Homo sapi
48	52	3.1	142323	9	AC044839	AC010988 Homo sapi
C	52	3.1	143125	9	AC092338	AC044839 Homo sapi
49	52	3.1	143125	9	AC092338	AC092338 Homo sapi
50	52	3.1	143577	9	AC013828	AC092338 Homo sapi
C	52	3.1	146169	2	AC128680	AC013828 Homo sapi
51	52	3.1	155185	2	AC074374	AC128680 Homo sapi
52	52	3.1	168781	2	AC036164	AC074374 Homo sapi
53	52	3.1	178524	9	AC055120	AC036164 Homo sapi
C	52	3.1	178524	9	AC055120	AC055120 Homo sapi
54	52	3.1	178834	9	CNS07EFF	AC117448 Homo sapi
55	52	3.1	181755	2	AC026996	AC117448 Homo sapi
C	52	3.1	183438	9	AC007738	AC026996 Homo sapi
57	52	3.1	183438	9	AC007738	AC007738 Homo sapi
58	52	3.1	185402	9	AC006536	AC007738 Homo sapi
59	52	3.1	186643	2	AC117273	AC006536 Homo sapi
60	52	3.1	189615	9	AC118584	AC117273 Homo sapi
61	52	3.1	190907	2	AC114732	AC118584 Homo sapi
C	52	3.1	200292	2	AC087442	AC114732 Homo sapi
62	52	3.1	200292	2	AC087442	AC087442 Homo sapi
C	52	3.1	202058	2	AC016913	AC087442 Homo sapi
63	52	3.1	214160	9	CNS00YVH	AC016913 Homo sapi
C	52	3.1	214160	9	CNS00YVH	AC094021 Homo sapi
64	52	3.1	228097	9	CNS00YVH	AC094021 Homo sapi
C	52	3.1	228097	9	CNS00YVH	AC094021 Homo sapi
65	52	3.1	228097	9	CNS00YVH	AC094021 Homo sapi

66	52	3.1	229155	9	AC018755	Homo sapi	139	50	3.0	179953	9	AC011475	Homo sapi
67	51	3.0	43711	9	AC079449	Homo sapi	140	50	3.0	181688	2	AP005354	Homo sapi
68	51	3.0	44350	2	AC105099	Homo sapi	c 141	50	3.0	185916	2	AP001568	Homo sapi
c 69	51	3.0	44350	2	AC105099	Homo sapi	c 142	50	3.0	186897	9	AC063932	Homo sapi
70	51	3.0	73688	9	AP002343	Homo sapi	c 143	50	3.0	186984	2	AC109496	Homo sapi
71	51	3.0	85311	2	AL354912	Homo sapi	c 144	50	3.0	188076	2	AC023629	Homo sapi
72	51	3.0	90125	9	AL353708	Homo sapi	c 145	50	3.0	188779	9	AC090701	Homo sapi
73	51	3.0	123253	9	AC004971	Homo sapi	c 146	50	3.0	189864	9	AC091719	Homo sapi
74	51	3.0	143364	2	AC004886	Homo sapi	c 147	50	3.0	194020	9	AC096679	Homo sapi
75	51	3.0	151348	2	AL138794	Homo sapi	c 148	50	3.0	194296	9	AL354864	Homo sapi
c 76	51	3.0	152528	2	AC015807	Homo sapi	c 149	50	3.0	194372	9	AC019184	Homo sapi
77	51	3.0	153042	9	AC016063	Homo sapi	c 150	50	3.0	196697	2	AL390062	Homo sapi
78	51	3.0	153496	2	AP002984	Homo sapi	c 151	50	3.0	196801	2	AC127390	Homo sapi
79	51	3.0	164472	9	AC018897	Homo sapi	c 152	50	3.0	197236	2	AC018710	Homo sapi
80	51	3.0	165044	2	AC007547	Homo sapi	c 153	50	3.0	197702	2	AC087178	Homo sapi
c 81	51	3.0	172321	9	AC010963	Homo sapi	c 154	50	3.0	199768	2	AP001589	Homo sapi
c 82	51	3.0	172321	9	AC010963	Homo sapi	c 155	50	3.0	199912	9	AC008430	Homo sapi
c 83	51	3.0	177720	2	AC019193	Homo sapi	c 156	50	3.0	212237	9	AC079235	Homo sapi
84	51	3.0	184295	2	AC021138	Homo sapi	c 157	50	3.0	224271	9	AC026803	Homo sapi
85	51	3.0	191659	9	AC069381	Homo sapi	c 158	50	3.0	225208	2	AC015561	Homo sapi
86	51	3.0	198565	2	AC016750	Homo sapi	c 159	50	3.0	226378	2	AC024583	Homo sapi
87	51	3.0	206454	9	AC018737	Homo sapi	c 160	50	3.0	266738	2	AL138896	Homo sapi
c 88	51	3.0	214017	2	AC026359	Homo sapi	c 161	49	2.9	301008	9	AC104802	Homo sapi
89	51	3.0	248281	9	AC008737	Homo sapi	c 162	49	2.9	336660	9	HSJ300013	Human DNA
90	50	3.0	18106	9	AL450407	Homo sapi	c 163	49	2.9	74516	2	AC025641	Homo sapi
91	50	3.0	37402	9	AC009002	Homo sapi	c 164	49	2.9	75582	9	AL138729	Human DNA
92	50	3.0	39531	9	AF024534	Homo sapi	c 165	49	2.9	77702	9	AL138743	Human DNA
c 93	50	3.0	40229	9	AC074139	Homo sapi	c 166	49	2.9	85718	9	AC008438	Homo sapi
c 94	50	3.0	43699	9	AC005934	Homo sapi	c 167	49	2.9	86825	9	AC006967	Homo sapi
c 95	50	3.0	44226	9	AY130859	Homo sapi	c 168	49	2.9	90543	9	AC090287	Homo sapi
96	50	3.0	51129	9	AL158013	Human DNA	c 169	49	2.9	126111	2	AL606518	Homo sapi
c 97	50	3.0	59522	2	AL513189	Human DNA	c 170	49	2.9	129293	9	HS1172N10	Human DNA
c 98	50	3.0	67521	2	AC103845	Homo sapi	c 171	49	2.9	134341	9	AC099794	Homo sapi
c 99	50	3.0	69035	2	AC119727	Continuation (4 of	c 172	49	2.9	146420	2	AP001140	Homo sapi
100	50	3.0	78956	2	AC022808	Homo sapi	c 173	49	2.9	150918	2	AC027364	Homo sapi
c 101	50	3.0	84912	9	AF024533	Homo sapi	c 174	49	2.9	151639	2	AP001193	Homo sapi
102	50	3.0	99128	2	AC007948	Genomic s	c 175	49	2.9	159439	9	AC074008	Homo sapi
103	50	3.0	111862	2	AL139143	Homo sapi	c 176	49	2.9	160341	9	AL161629	Homo sapi
104	50	3.0	112831	9	AC004974	Homo sapi	c 177	49	2.9	164866	2	AC091102	Human DNA
105	50	3.0	122228	2	AC000003	Homo sapi	c 178	49	2.9	165351	9	AC019109	Homo sapi
c 106	50	3.0	123304	2	AF172274	Homo sapi	c 179	49	2.9	169616	9	AL355336	Human DNA
107	50	3.0	124497	9	AC120053	Homo sapi	c 180	49	2.9	172008	9	AC079602	Homo sapi
108	50	3.0	127140	9	AL136137	Homo sapi	c 181	49	2.9	172453	2	AC015889	Homo sapi
c 109	50	3.0	128736	9	AC011453	Homo sapi	c 182	49	2.9	175323	2	AC012168	Homo sapi
c 110	50	3.0	129586	9	AC005086	Homo sapi	c 183	49	2.9	182452	9	AL161731	Human DNA
c 111	50	3.0	134792	9	AC020942	Homo sapi	c 184	49	2.9	183461	9	AC006121	Homo sapi
112	50	3.0	135698	2	AL160287	Homo sapi	c 185	49	2.9	183515	9	AC107223	Homo sapi
113	50	3.0	139072	9	AL365218	Homo sapi	c 186	49	2.9	184057	9	AP000753	Homo sapi
c 114	50	3.0	142728	9	AL589943	Human DNA	c 187	49	2.9	184059	2	AC090382	Homo sapi
115	50	3.0	143686	2	AL049636	Human DNA	c 188	49	2.9	193104	2	AC090339	Homo sapi
c 116	50	3.0	14666	9	AC068836	Homo sapi	c 189	49	2.9	199134	2	AC023458	Homo sapi
117	50	3.0	145481	9	AL160287	Homo sapi	c 190	49	2.9	206019	2	AC078875	Homo sapi
c 118	50	3.0	145481	9	AL139801	Human DNA	c 191	49	2.9	221187	2	HS92N15	Homo sapi
119	50	3.0	146596	9	AC007631	Genomic s	c 192	48	2.9	1549	9	AF294262	Gorilla g
120	50	3.0	151182	9	AL611942	Human DNA	c 193	48	2.9	1560	9	AF294261	Pan panis
121	50	3.0	154547	2	AC016112	Homo sapi	c 194	48	2.9	1909	9	AF294261	Pan panis
c 122	50	3.0	159707	2	AL365403	Homo sapi	c 195	48	2.9	2880	9	AF375643	Pongo pyg
123	50	3.0	163360	2	AC091716	Homo sapi	c 196	48	2.9	2880	9	AF375643	Pongo pyg
c 124	50	3.0	164275	2	AC021634	Homo sapi	c 197	48	2.9	24132	9	D86992	Homo sapi
125	50	3.0	166702	9	AC012594	Homo sapi	c 198	48	2.9	29010	2	AC027293	Homo sapi
c 126	50	3.0	168000	2	AC062030	Homo sapi	c 199	48	2.9	35492	9	AC006125	Homo sapi
c 127	50	3.0	168659	2	AL359883	Homo sapi	c 200	48	2.9	37027	9	AL135844	Human DNA
128	50	3.0	168963	2	AC009705	Homo sapi	c 201	48	2.9	39019	9	D88270	Human DNA
c 129	50	3.0	169662	9	AC068944	Homo sapi	c 202	48	2.9	42442	2	AC103812	Homo sapi
130	50	3.0	171166	9	AC012354	Homo sapi	c 203	48	2.9	54171	9	AC118139	Homo sapi
c 131	50	3.0	173967	2	AC073343	Homo sapi	c 204	48	2.9	57573	2	AC087509	Homo sapi
c 132	50	3.0	174527	2	AC092859	Homo sapi	c 205	48	2.9	58072	2	AC126391	Homo sapi
133	50	3.0	175497	2	AL645737	Homo sapi	c 206	48	2.9	60538	2	AC102939	Homo sapi
c 134	50	3.0	175870	9	AC007637	Homo sapi	c 207	48	2.9	61798	9	AL512648	Human DNA
c 135	50	3.0	176783	9	AL139281	Human DNA	c 208	48	2.9	62440	9	HSJ7636L22	Human DNA
136	50	3.0	178218	2	AC016841	Homo sapi	c 209	48	2.9	64650	2	AC068864	Homo sapi
c 137	50	3.0	178736	9	AC103853	Homo sapi	c 210	48	2.9	64900	2	AC129497	Homo sapi
c 138	50	3.0	179394	9	AC010619	Homo sapi	c 211	48	2.9	64980	2	AC127539	Homo sapi

212	48	2.9	66624	2	AC024376	AC024376	Homo sapi	285	48	2.9	143860	9	AC009155	AC009155	Homo sapi
c 213	48	2.9	67400	2	AC048341	AC048341	Homo sapi	286	48	2.9	145253	9	AC006329	AC006329	Homo sapi
c 214	48	2.9	67565	2	AC126328	AC126328	Homo sapi	c 287	48	2.9	145962	9	AC098690	AC098690	Homo sapi
c 215	48	2.9	68001	9	AP001442	AP001442	Homo sapi	288	48	2.9	145988	2	AC024059	AC024059	Homo sapi
c 216	48	2.9	70065	9	AC012630	AC012630	Homo sapi	289	48	2.9	146167	9	AC087588	AC087588	Homo sapi
c 217	48	2.9	71527	2	AC131568	AC131568	Homo sapi	290	48	2.9	147109	9	AL358779	AL358779	Human DNA
c 218	48	2.9	72333	2	AL512329	AL512329	Human DNA	c 291	48	2.9	147277	2	AC099811	AC099811	Homo sapi
c 219	48	2.9	72826	9	AL391539	AL391539	Human DNA	c 292	48	2.9	147418	9	AC023855	AC023855	Homo sapi
c 220	48	2.9	73304	2	AC016571	AC016571	Homo sapi	c 293	48	2.9	148116	2	AC078838	AC078838	Homo sapi
c 221	48	2.9	74356	2	AC016511	AC016511	Homo sapi	c 294	48	2.9	148295	2	AC090510	AC090510	Homo sapi
c 222	48	2.9	75663	9	AC068987	AC068987	Homo sapi	295	48	2.9	148401	9	AC092118	AC092118	Homo sapi
c 223	48	2.9	76702	9	HS756P4	AL035072	Human DNA	c 296	48	2.9	149538	9	AC002351	AC002351	Homo sapi
c 224	48	2.9	82781	9	AC103691	AC103691	Homo sapi	c 297	48	2.9	149668	9	AC025268	AC025268	Homo sapi
c 225	48	2.9	83057	9	AL445703	AL445703	Human DNA	c 298	48	2.9	149790	2	AC022106	AC022106	Homo sapi
c 226	48	2.9	83412	2	AC106778	AC106778	Homo sapi	c 299	48	2.9	150150	2	AC022003	AC022003	Homo sapi
c 227	48	2.9	84087	9	AC010492	AC010492	Homo sapi	c 300	48	2.9	150913	9	AC083875	AC083875	Homo sapi
c 228	48	2.9	84129	9	AC103813	AC103813	Homo sapi	301	48	2.9	151289	9	AL158164	AL158164	Human DNA
c 229	48	2.9	87216	2	AC016376	AC016376	Homo sapi	302	48	2.9	151432	9	AL732579	AL732579	Human DNA
c 230	48	2.9	89340	2	AL355597	AL355597	Human DNA	c 303	48	2.9	151770	9	AC002066	AC002066	Homo sapi
c 231	48	2.9	90968	2	AC108722	AC108722	Homo sapi	c 304	48	2.9	151865	2	AL356236	AL356236	Homo sapi
c 232	48	2.9	96281	9	AP000659	AP000659	Homo sapi	305	48	2.9	152058	30	AC025232	AC025232	Homo sapi
c 233	48	2.9	96398	9	AL591367	AL591367	Human DNA	306	48	2.9	152081	9	AC007381	AC007381	Homo sapi
c 234	48	2.9	97137	2	AL355172	AL355172	Homo sapi	307	48	2.9	153078	9	AC022493	AC022493	Homo sapi
c 235	48	2.9	98357	2	AL391871	AL391871	Homo sapi	c 308	48	2.9	153120	9	AC019050	AC019050	Homo sapi
c 236	48	2.9	98804	2	AC106723_3	Continuaton (4 of	c 309	48	2.9	153553	2	AL513012	AL513012	Homo sapi	
c 237	48	2.9	100000	9	AP000141	AP000141	Homo sapi	310	48	2.9	153951	2	AC019027	AC019027	Homo sapi
c 238	48	2.9	100079	9	AC108002	AC108002	Homo sapi	c 311	48	2.9	154036	9	AC008068	AC008068	Homo sapi
c 239	48	2.9	100760	9	AC105423	AC105423	Homo sapi	312	48	2.9	154831	2	AC019085	AC019085	Homo sapi
c 240	48	2.9	103247	2	AE176678	AE176678	Homo sapi	313	48	2.9	154959	2	AC009282	AC009282	Homo sapi
c 241	48	2.9	104198	9	AL136120	AL136120	Human DNA	c 314	48	2.9	154983	9	AC018866	AC018866	Homo sapi
c 242	48	2.9	104794	9	AC104046	AC104046	Homo sapi	c 315	48	2.9	155257	2	AC126336	AC126336	Homo sapi
c 243	48	2.9	105087	9	AC093884	AC093884	Homo sapi	316	48	2.9	155375	2	AC098854	AC098854	Homo sapi
c 244	48	2.9	105156	2	AF235094	AF235094	Homo sapi	317	48	2.9	155523	9	AP001528	AP001528	Homo sapi
c 245	48	2.9	108503	9	AL627424	AL627424	Human DNA	c 318	48	2.9	155848	9	AC104996	AC104996	Homo sapi
c 246	48	2.9	109149	2	AP000681	AP000681	Homo sapi	319	48	2.9	155897	2	AP000823	AP000823	Homo sapi
c 247	48	2.9	109878	9	AL390208	AL390208	Human DNA	320	48	2.9	155935	2	AC006442	AC006442	Homo sapi
c 248	48	2.9	110000	2	AC117465_3	Continuaton (4 of	c 321	48	2.9	156601	9	HS212G6	HS212G6	Homo sapi	
c 249	48	2.9	111421	9	AC010632	AC010632	Homo sapi	322	48	2.9	156784	2	AC016390	AC016390	Homo sapi
c 250	48	2.9	111469	2	AC022470	AC022470	Homo sapi	c 323	48	2.9	157028	9	AL139000	AL139000	Human DNA
c 251	48	2.9	111746	9	HS327J16	AL008583	Human DNA	324	48	2.9	157392	2	AC129984	AC129984	Homo sapi
c 252	48	2.9	111998	9	AC011509	AC011509	Homo sapi	325	48	2.9	157959	9	AP003101	AP003101	Homo sapi
c 253	48	2.9	112291	2	AL137799	AL137799	Homo sapi	326	48	2.9	158419	2	AC055855	AC055855	Homo sapi
c 254	48	2.9	113761	2	AL355494	AL355494	Homo sapi	327	48	2.9	159621	9	AC008301	AC008301	Homo sapi
c 255	48	2.9	115932	9	HS110I7	AL031589	Human DNA	328	48	2.9	159807	2	AP002389	AP002389	Homo sapi
c 256	48	2.9	119420	9	AC027485	AC027485	Homo sapi	329	48	2.9	160196	2	AC016701	AC016701	Homo sapi
c 257	48	2.9	120797	2	AC087512	AC087512	Pan trogl	330	48	2.9	160285	2	AC016668	AC016668	Homo sapi
c 258	48	2.9	121129	9	AL357315	AL357315	Human DNA	331	48	2.9	160421	2	AC025005	AC025005	Homo sapi
c 259	48	2.9	121394	2	AL137799	AL137799	Homo sapi	c 332	48	2.9	161014	9	AP000088	AP000088	Homo sapi
c 260	48	2.9	122487	9	AC010270	AC010270	Homo sapi	c 333	48	2.9	161039	2	AC011276	AC011276	Homo sapi
c 261	48	2.9	123284	2	HS0998N21	AL109948	Homo sapi	c 334	48	2.9	161226	9	AP001619	AP001619	Homo sapi
c 262	48	2.9	126613	9	HSJ603J14	AL122001	Human DNA	c 335	48	2.9	161586	9	AC090179	AC090179	Homo sapi
c 263	48	2.9	126736	2	AC044788	AC044788	Homo sapi	336	48	2.9	162163	2	AC021635	AC021635	Homo sapi
c 264	48	2.9	127078	9	AP000708	AP000708	Homo sapi	c 337	48	2.9	162354	9	AF260225	AF260225	Homo sapi
c 265	48	2.9	129517	2	AC016400	AC016400	Homo sapi	c 338	48	2.9	162445	9	AL158151	AL158151	Human DNA
c 266	48	2.9	129801	9	AC073614	AC073614	Homo sapi	c 339	48	2.9	162862	2	AC021790	AC021790	Homo sapi
c 267	48	2.9	129819	9	AC104049	AC104049	Homo sapi	c 340	48	2.9	163022	2	AC011802	AC011802	Homo sapi
c 268	48	2.9	129984	9	AC115620	AC115620	Homo sapi	341	48	2.9	163024	9	AP000758	AP000758	Homo sapi
c 269	48	2.9	131303	9	AL355974	AL355974	Human DNA	342	48	2.9	163046	9	AP000889	AP000889	Homo sapi
c 270	48	2.9	132486	9	HSF10178	AL035367	Homo sapi	c 343	48	2.9	163114	9	AC096886	AC096886	Homo sapi
c 271	48	2.9	133104	9	AL451070	AL451070	Human DNA	c 344	48	2.9	163167	9	AC099784	AC099784	Homo sapi
c 272	48	2.9	135805	2	AC069511	AC069511	Homo sapi	c 345	48	2.9	163303	2	AC104467	AC104467	Homo sapi
c 273	48	2.9	135875	9	AC109445	AC109445	Homo sapi	c 346	48	2.9	163600	9	AC093563	AC093563	Homo sapi
c 274	48	2.9	136000	9	AP003119	AP003119	Homo sapi	347	48	2.9	164198	2	AL390796	AL390796	Homo sapi
c 275	48	2.9	136385	9	AC006006	AC006006	Homo sapi	348	48	2.9	164958	9	AC099340	AC099340	Homo sapi
c 276	48	2.9	138876	2	AL590067	AL590067	Homo sapi	c 349	48	2.9	164963	2	AL446025	AL446025	Homo sapi
c 277	48	2.9	138890	9	AC007970	AC007970	Homo sapi	c 350	48	2.9	165106	2	AC016034	AC016034	Homo sapi
c 278	48	2.9	138928	9	AC064847	AC064847	Homo sapi	c 351	48	2.9	165571	2	AC124263	AC124263	Homo sapi
c 279	48	2.9	139247	2	AC127600	AC127600	Homo sapi	352	48	2.9	165719	2	AC127521	AC127521	Homo sapi
c 280	48	2.9	140092	9	AL162615	AL162615	Human DNA	c 353	48	2.9	165978	9	AC013421	AC013421	Homo sapi
c 281	48	2.9	140546	2	AC122688	AC122688	Homo sapi	354	48	2.9	166439	2	AC013035	AC013035	Homo sapi
c 282	48	2.9	140712	2	AC022223	AC022223	Homo sapi	355	48	2.9	166873	9	AC008554	AC008554	Homo sapi
c 283	48	2.9	141277	9	AL357045	AL357045	Human DNA	c 356	48	2.9	167408	30	AC023020	AC023020	Homo sapi
c 284	48	2.9	143133	2	AC021655	AC021655	Human sapi	357	48	2.9	167830	9	AL136219	AL136219	Human DNA

C 358	48	2.9 167886	2	AP002405	AP002405 Homo sapi	431	48	2.9 194575	2	AC023140	AC023140 Homo sapi
C 359	48	2.9 167924	2	AC067891	AC067891 Homo sapi	432	48	2.9 194624	9	AC008742	AC008742 Homo sapi
C 360	48	2.9 168063	2	AC090699	AC090699 Homo sapi	C 433	48	2.9 194880	2	AC091929	AC091929 Homo sapi
C 361	48	2.9 168174	2	AC046169	AC046169 Homo sapi	C 434	48	2.9 195364	9	HS431A14	285996 Human DNA s
C 362	48	2.9 168624	9	CNS01D7N	AL132801 Human chr	C 435	48	2.9 195840	9	CNS0000P	AL049869 Human chr
C 363	48	2.9 168839	2	AC022826	AC022826 Homo sapi	C 436	48	2.9 196421	2	AC078837	AC078837 Homo sapi
C 364	48	2.9 169118	2	AC096645	AC096645 Homo sapi	C 437	48	2.9 198185	2	AC087691	AC087691 Homo sapi
C 365	48	2.9 169352	2	AL592211	AL592211 Homo sapi	C 438	48	2.9 199268	2	AC099523	AC099523 Homo sapi
C 366	48	2.9 169476	9	CNS01D7W	AL136419 Human chr	C 439	48	2.9 199297	2	AC103975	AC103975 Homo sapi
C 367	48	2.9 169586	9	AC078883	AC078883 Homo sapi	C 440	48	2.9 199721	9	AC092636	AC092636 Homo sapi
C 368	48	2.9 170200	9	AC008012	AC008012 Homo sapi	C 441	48	2.9 199959	2	AC093719	AC093719 Homo sapi
C 369	48	2.9 170470	2	AC022696	AC022696 Homo sapi	C 442	48	2.9 200000	2	AC006294	AC006294 Homo sapi
C 370	48	2.9 171976	9	AP003069	AP003069 Homo sapi	C 443	48	2.9 202398	2	AC062014	AC062014 Homo sapi
C 371	48	2.9 172738	2	AC074104	AC074104 Homo sapi	C 444	48	2.9 203269	2	AC080187	AC080187 Homo sapi
C 372	48	2.9 172753	2	AC032019	AC032019 Homo sapi	C 445	48	2.9 203748	2	AC069250	AC069250 Homo sapi
C 373	48	2.9 173040	9	AC012442	AC012442 Homo sapi	C 446	48	2.9 203788	2	AC040970	AC040970 Homo sapi
C 374	48	2.9 173073	9	AC022817	AC022817 Homo sapi	C 447	48	2.9 204268	2	AC092163	AC092163 Homo sapi
C 375	48	2.9 173256	2	AC073301	AC073301 Homo sapi	C 448	48	2.9 205819	2	AC106805	AC106805 Homo sapi
C 376	48	2.9 173354	2	AC027554	AC027554 Homo sapi	C 449	48	2.9 205919	2	AC012279	AC012279 Homo sapi
C 377	48	2.9 173364	2	AC009807	AC009807 Homo sapi	C 450	48	2.9 208065	9	AC008676	AC008676 Homo sapi
C 378	48	2.9 173509	2	AC008557	AC008557 Homo sapi	C 451	48	2.9 209216	9	AF190464	AF190464 Homo sapi
C 379	48	2.9 173889	2	AC013691	AC013691 Homo sapi	C 452	48	2.9 209729	2	AC094014	AC094014 Papio cyn
C 380	48	2.9 174128	9	AL137139	AL137139 Human DNA	C 453	48	2.9 209861	2	AC021196	AC021196 Homo sapi
C 381	48	2.9 174525	2	AC067763	AC067763 Homo sapi	C 454	48	2.9 211748	9	AC010146	AC010146 Homo sapi
C 382	48	2.9 174673	9	AC019187	AC019187 Homo sapi	C 455	48	2.9 21275	9	AC012076	AC012076 Homo sapi
C 383	48	2.9 175625	2	AC068475	AC068475 Homo sapi	C 456	48	2.9 214565	2	AL357493	AL357493 Homo sapi
C 384	48	2.9 175808	9	AC068727	AC068727 Homo sapi	C 457	48	2.9 215111	9	AC010255	AC010255 Homo sapi
C 385	48	2.9 175989	2	AL590962	AL590962 Homo sapi	C 458	48	2.9 215385	9	AF216808	AF216808 Homo sapi
C 386	48	2.9 176380	2	AC011371	AC011371 Homo sapi	C 459	48	2.9 216252	2	AC009064	AC009064 Homo sapi
C 387	48	2.9 177262	2	AC123979	AC123979 Homo sapi	C 460	48	2.9 216260	9	AC099518	AC099518 Homo sapi
C 388	48	2.9 177277	9	AL353668	AL353668 Human DNA	C 461	48	2.9 216504	2	AC090268	AC090268 Homo sapi
C 389	48	2.9 177289	2	AC090176	AC090176 Homo sapi	C 462	48	2.9 217460	9	AC040162	AC040162 Homo sapi
C 390	48	2.9 177409	2	AC009128	AC009128 Homo sapi	C 463	48	2.9 218355	2	AC027243	AC027243 Homo sapi
C 391	48	2.9 178152	9	AC096633	AC096633 Homo sapi	C 464	48	2.9 221475	9	AC009562	AC009562 Homo sapi
C 392	48	2.9 178189	9	AP001823	AP001823 Homo sapi	C 465	48	2.9 227137	9	AC098481	AC098481 Homo sapi
C 393	48	2.9 178820	9	AC020552	AC020552 Homo sapi	C 466	48	2.9 231260	2	AF215845	AF215845 Homo sapi
C 394	48	2.9 178985	9	HS340B19	AL033519 Human DNA	C 467	48	2.9 238540	2	AL160172	AL160172 Homo sapi
C 395	48	2.9 179393	9	AC010636	AC010636 Homo sapi	C 468	48	2.9 243230	9	AC008741	AC008741 Homo sapi
C 396	48	2.9 179871	2	AC024602	AC024602 Homo sapi	C 469	48	2.9 250529	9	HUAE000658	AE000658 Homo sapi
C 397	48	2.9 180253	2	AL513362	AL513362 Homo sapi	C 470	48	2.9 262951	2	AC124848	AC124848 Homo sapi
C 398	48	2.9 180277	2	AC103559	AC103559 Homo sapi	C 471	48	2.9 283396	2	AC093024	AC093024 Homo sapi
C 399	48	2.9 180317	2	AC025590	AC025590 Homo sapi	C 472	48	2.9 301692	9	D87675	D87675 Homo sapien
C 400	48	2.9 181594	2	AC019167	AC019167 Homo sapi	C 473	48	2.9 316296	9	AC092285	AC092285 Homo sapi
C 401	48	2.9 181991	2	AC068322	AC068322 Homo sapi	C 474	48	2.9 340000	9	AP001695	AP001695 Homo sapi
C 402	48	2.9 182632	9	AC007282	AC007282 Homo sapi	C 475	47	2.8 729	9	HS4343724	AJ343724 Homo sapi
C 403	48	2.9 182959	9	AL136982	AL136982 Homo sapi	C 476	47	2.8 976	9	HS4343559	AJ343559 Homo sapi
C 404	48	2.9 183443	2	AC068145	AC068145 Homo sapi	C 477	47	2.8 45848	9	AL133385	AL133385 Human DNA
C 405	48	2.9 184439	2	AP003085	AP003085 Homo sapi	C 478	47	2.8 72147	2	AC021693	AC021693 Homo sapi
C 406	48	2.9 184555	2	AC079379	AC079379 Homo sapi	C 479	47	2.8 87100	9	HS43814	Z97635 Human DNA s
C 407	48	2.9 185351	9	AC015801	AC015801 Homo sapi	C 480	47	2.8 92312	9	AC112495	AC112495 Homo sapi
C 408	48	2.9 185378	9	AC118754	AC118754 Homo sapi	C 481	47	2.8 104531	9	HS43303A1	AL096800 Human DNA
C 409	48	2.9 185666	2	AC048389	AC048389 Homo sapi	C 482	47	2.8 115383	9	AC026716	AC026716 Homo sapi
C 410	48	2.9 185735	9	AC078785	AC078785 Homo sapi	C 483	47	2.8 118607	2	AC114484	AC114484 Homo sapi
C 411	48	2.9 186304	9	AC090136	AC090136 Homo sapi	C 484	47	2.8 118695	9	HS7672M15	AL049643 Human DNA
C 412	48	2.9 186625	9	AC006487	AC006487 Homo sapi	C 485	47	2.8 123110	9	AC005383	AC005383 Homo sapi
C 413	48	2.9 187380	9	AC023946	AC023946 Homo sapi	C 486	47	2.8 123953	2	AP001259	AP001259 Homo sapi
C 414	48	2.9 188107	9	CNS01DWE	AL137129 Human chr	C 487	47	2.8 131465	9	AL731573	AL731573 Human DNA
C 415	48	2.9 188285	2	AC079451	AC079451 Homo sapi	C 488	47	2.8 131944	9	AC097659	AC097659 Homo sapi
C 416	48	2.9 188948	2	AC005832	AC005832 Homo sapi	C 489	47	2.8 132887	9	AL160408	AL160408 Human DNA
C 417	48	2.9 190225	2	AC011191	AC011191 Homo sapi	C 490	47	2.8 133225	9	AC108089	AC108089 Homo sapi
C 418	48	2.9 190466	2	AC022285	AC022285 Homo sapi	C 491	47	2.8 134841	9	AC018840	AC018840 Homo sapi
C 419	48	2.9 190870	2	AL157894	AL157894 Homo sapi	C 492	47	2.8 140334	9	AC021802	AC021802 Homo sapi
C 420	48	2.9 190815	9	AC007066	AC007066 Homo sapi	C 493	47	2.8 143065	9	HS20208	AL031848 Human DNA
C 421	48	2.9 190824	2	AL731645	AL731645 Homo sapi	C 494	47	2.8 152245	2	AC119033	AC119033 Homo sapi
C 422	48	2.9 190941	2	AC108045	AC108045 Homo sapi	C 495	47	2.8 153311	2	AC022829	AC022829 Homo sapi
C 423	48	2.9 191236	2	AC068942	AC068942 Homo sapi	C 496	47	2.8 153311	2	AC107297	AC107297 Homo sapi
C 424	48	2.9 191240	2	AC027738	AC027738 Homo sapi	C 497	47	2.8 155523	9	AP001528	AP001528 Homo sapi
C 425	48	2.9 191365	2	AC020588	AC020588 Homo sapi	C 498	47	2.8 157382	2	AC009279	AC009279 Homo sapi
C 426	48	2.9 191525	9	AC103830	AC103830 Homo sapi	C 499	47	2.8 163592	9	AC110048	AC110048 Homo sapi
C 427	48	2.9 191594	9	AL160282	AL160282 Human DNA	C 500	47	2.8 163690	2	AC083781	AC083781 Homo sapi
C 428	48	2.9 193099	9	AC092140	AC092140 Homo sapi	C 501	47	2.8 165245	2	AC027599	AC027599 Homo sapi
C 429	48	2.9 193709	9	AL390838	AL390838 Human DNA	C 502	47	2.8 165785	9	AC113223	AC113223 Homo sapi
C 430	48	2.9 194237	9	AC068870	AC068870 Homo sapi	C 503	47				



c 650	44	2.6	59919	2	AC124648	Homo sapi	c 723	44	2.6	169963	9	AL158207	Human DNA
c 651	44	2.6	62000	9	AP001984	Homo sapi	c 724	44	2.6	172024	2	AC078971	Homo sapi
c 652	44	2.6	67024	9	AP001494	Homo sapi	c 725	44	2.6	172862	2	AL645796	Homo sapi
c 653	44	2.6	70193	9	AL627303	Human DNA	c 726	44	2.6	174441	2	AC009683	Homo sapi
c 654	44	2.6	73650	2	AC027703	Homo sapi	c 727	44	2.6	175522	2	AC023970	Homo sapi
c 655	44	2.6	80723	9	AC113348	Homo sapi	c 728	44	2.6	176162	9	AC093376	Homo sapi
c 656	44	2.6	82938	9	AC079250	Homo sapi	c 729	44	2.6	176398	9	AC092032	Homo sapi
c 657	44	2.6	83714	9	AL136125	Human DNA	c 730	44	2.6	176498	9	AC003035	Homo sapi
c 658	44	2.6	86654	9	HSJ22117	Human DNA	c 731	44	2.6	176930	2	AC090010	Homo sapi
c 659	44	2.6	87244	9	AL356741	Human DNA	c 732	44	2.6	177344	9	AL356953	Human DNA
c 660	44	2.6	94023	9	AC008720	Homo sapi	c 733	44	2.6	177594	9	AC026368	Homo sapi
c 661	44	2.6	94893	9	AC064860	Homo sapi	c 734	44	2.6	178157	2	AC117404	Homo sapi
c 662	44	2.6	95214	9	AC012457	Homo sapi	c 735	44	2.6	179343	9	AC007281	Homo sapi
c 663	44	2.6	100733	9	HS201D7	Human DNA	c 736	44	2.6	180446	2	AP001927	Homo sapi
c 664	44	2.6	102517	2	AP002876	Homo sapi	c 737	44	2.6	180615	2	AC025549	Homo sapi
c 665	44	2.6	110184	9	AC020906	Homo sapi	c 738	44	2.6	181118	9	AC067871	Homo sapi
c 666	44	2.6	116814	9	AC096660	Homo sapi	c 739	44	2.6	181135	2	AP001016	Homo sapi
c 667	44	2.6	117711	9	AP000662	Homo sapi	c 740	44	2.6	181259	9	AC063951	Homo sapi
c 668	44	2.6	118103	2	AC009817	Homo sapi	c 741	44	2.6	181663	9	AC092415	Homo sapi
c 669	44	2.6	118958	9	AC092579	Homo sapi	c 742	44	2.6	181991	9	AC025882	Homo sapi
c 670	44	2.6	118981	9	AC092546	Homo sapi	c 743	44	2.6	183778	2	AC011004	Homo sapi
c 671	44	2.6	119794	9	HSJ295844	Human DNA	c 744	44	2.6	183935	9	AL359382	Human DNA
c 672	44	2.6	120766	9	AC004150	Homo sapi	c 745	44	2.6	183946	2	AL354719	Homo sapi
c 673	44	2.6	122186	6	AX053112	Sequence	c 746	44	2.6	184695	2	AC061711	Homo sapi
c 674	44	2.6	122186	9	AC004466	Homo sapi	c 747	44	2.6	185019	9	AC008732	Homo sapi
c 675	44	2.6	127687	2	AC005959	Homo sapi	c 748	44	2.6	185165	9	AC008732	Homo sapi
c 676	44	2.6	128978	9	AC002073	Human PAC	c 749	44	2.6	185408	2	AC021053	Homo sapi
c 677	44	2.6	133401	2	AC079257	Homo sapi	c 750	44	2.6	186048	2	AC074190	Homo sapi
c 678	44	2.6	134091	9	AC023349	Homo sapi	c 751	44	2.6	187728	2	AC062011	Homo sapi
c 679	44	2.6	136372	9	AC103828	Homo sapi	c 752	44	2.6	188868	9	AC023232	Homo sapi
c 680	44	2.6	137213	9	AC005343	Homo sapi	c 753	44	2.6	190054	2	AC114730	Homo sapi
c 681	44	2.6	137856	9	AC097720	Homo sapi	c 754	44	2.6	191152	9	AC005722	Homo sapi
c 682	44	2.6	141272	9	AP001976	Homo sapi	c 755	44	2.6	193608	2	AC019426	Homo sapi
c 683	44	2.6	141273	9	AL611933	Human DNA	c 756	44	2.6	193887	9	AF313041	Homo sapi
c 684	44	2.6	143132	2	AC027443	Homo sapi	c 757	44	2.6	193957	2	AC090198	Homo sapi
c 685	44	2.6	147124	2	AC005913	Homo sapi	c 758	44	2.6	194109	2	AC121338	Homo sapi
c 686	44	2.6	147694	2	AC090628	Homo sapi	c 759	44	2.6	194171	2	AC097350	Homo sapi
c 687	44	2.6	147803	9	AL356867	Human DNA	c 760	44	2.6	194520	9	AL356915	Human DNA
c 688	44	2.6	147913	9	HSJ613A2	Human DNA	c 761	44	2.6	196927	2	HS1091N2	Homo sapi
c 689	44	2.6	148708	9	AC079953	Homo sapi	c 762	44	2.6	197175	9	AC110292	Homo sapi
c 690	44	2.6	148811	2	AC026567	Homo sapi	c 763	44	2.6	199240	2	AC010287	Homo sapi
c 691	44	2.6	148970	9	AC105268	Homo sapi	c 764	44	2.6	199287	2	AC068659	Homo sapi
c 692	44	2.6	150336	9	HSJ395C13	Homo sapi	c 765	44	2.6	199347	2	AC091575	Homo sapi
c 693	44	2.6	150681	9	AC006011	Human DNA	c 766	44	2.6	199392	2	AC021522	Homo sapi
c 694	44	2.6	151640	2	AC015816	Homo sapi	c 767	44	2.6	200791	2	AC125507	Homo sapi
c 695	44	2.6	152118	2	AC013263	Homo sapi	c 768	44	2.6	203868	9	AP002381	Paplo cyn
c 696	44	2.6	152212	9	AC106736	Homo sapi	c 769	44	2.6	204129	2	AC026293	Homo sapi
c 697	44	2.6	152824	9	AC107072	Homo sapi	c 770	44	2.6	206223	2	AC121329	Homo sapi
c 698	44	2.6	152901	2	AC055781	Homo sapi	c 771	44	2.6	206413	9	CNS06C7V	Human chr
c 699	44	2.6	153248	9	AC112191	Homo sapi	c 772	44	2.6	206616	2	AC090622	Homo sapi
c 700	44	2.6	153860	2	AC016301	Homo sapi	c 773	44	2.6	206682	2	AC060790	Homo sapi
c 701	44	2.6	154140	9	CNS06C7P	Human chr	c 774	44	2.6	207062	9	AF295017	Homo sapi
c 702	44	2.6	154509	2	AC026075	Homo sapi	c 775	44	2.6	207069	9	AC090922	Homo sapi
c 703	44	2.6	154683	9	AC006157	Homo sapi	c 776	44	2.6	207411	2	AC009956	Homo sapi
c 704	44	2.6	154937	2	AP002382	Homo sapi	c 777	44	2.6	211740	2	AC109826	Homo sapi
c 705	44	2.6	156763	9	AC005901	Homo sapi	c 778	44	2.6	215399	9	AL160278	Human DNA
c 706	44	2.6	158083	9	AL162493	Human DNA	c 779	44	2.6	221673	2	AC012183	Homo sapi
c 707	44	2.6	158090	9	AC006353	Homo sapi	c 780	44	2.6	320902	2	AL161456	Homo sapi
c 708	44	2.6	158349	2	AC009438	Homo sapi	c 781	44	2.6	3239	9	AB048967	Homo sapi
c 709	44	2.6	159704	9	AC106000	Homo sapi	c 782	44	2.6	38448	9	AF527540	Macaca fa
c 710	44	2.6	159784	2	AC061996	Homo sapi	c 783	44	2.6	38448	9	AC000397	Genomic S
c 711	44	2.6	163811	2	AC108689	Homo sapi	c 784	44	2.6	41219	9	AC005382	Homo sapi
c 712	44	2.6	164396	2	AC016519	Homo sapi	c 785	44	2.6	42790	9	HSJ315767	Homo sapi
c 713	44	2.6	164508	2	AC104130	Homo sapi	c 786	44	2.6	47015	2	AC091191	Homo sapi
c 714	44	2.6	166889	2	AC027008	Homo sapi	c 787	44	2.6	54354	2	AC131046	Homo sapi
c 715	44	2.6	167633	9	AL513534	Human DNA	c 788	44	2.6	62900	2	AC016244	Homo sapi
c 716	44	2.6	168006	2	AC090699	Homo sapi	c 789	44	2.6	6343	2	AC131261	Homo sapi
c 717	44	2.6	168063	2	AC058819	Homo sapi	c 790	44	2.6	6343	2	AC131075	Mus muscu
c 718	44	2.6	168205	2	AC010624	Homo sapi	c 791	44	2.6	70432	2	AC124278	Homo sapi
c 719	44	2.6	168230	2	AC092678	Homo sapi	c 792	44	2.6	70432	2	AC124278	Homo sapi
c 720	44	2.6	168395	2	AC007649	Homo sapi	c 793	44	2.6	71319	2	AC123990	Homo sapi
c 721	44	2.6	168623	9	AC113557	Homo sapi	c 794	44	2.6	72231	2	AC124284	Homo sapi
c 722	44	2.6	169565	2	AC113557	Homo sapi	c 795	44	2.6	72231	2	AC124284	Homo sapi

c 796	43	2.6	78366	9	AL590398	AL590398	Human DNA	869	43	2.6	180948	9	AC114399	AC114399	Homo sapi
c 797	43	2.6	81483	9	AL354699	AL354699	Human DNA	870	43	2.6	181991	2	AC068322	AC068322	Homo sapi
c 798	43	2.6	84570	9	AL359552	AL359552	Human DNA	871	43	2.6	182547	9	AP001201	AP001201	Homo sapi
c 799	43	2.6	85149	9	AL137879	AL137879	Human DNA	872	43	2.6	184264	2	AL772267	AL772267	Homo sapi
800	43	2.6	87903	9	AL732364	AL732364	Human DNA	873	43	2.6	184513	2	AL158826	AL158826	Human DNA
801	43	2.6	87903	9	AL732364	AL732364	Human DNA	874	43	2.6	184513	2	AL158826	AL158826	Human DNA
802	43	2.6	97495	2	AC093014	AC093014	Homo sapi	c 874	43	2.6	186044	2	AC004831	AC004831	Homo sapi
803	43	2.6	98829	2	AC002345	AC002345	Homo sapi	875	43	2.6	186482	2	AC021071	AC021071	Homo sapi
c 804	43	2.6	104307	9	AL591845	AL591845	Human DNA	876	43	2.6	186626	9	AP000902	AP000902	Homo sapi
c 805	43	2.6	104420	2	AC113554	AC113554	Homo sapi	c 877	43	2.6	187414	2	AC087163	AC087163	Homo sapi
c 806	43	2.6	106172	9	AL591122	AL591122	Human DNA	878	43	2.6	188416	2	AC093855	AC093855	Homo sapi
c 807	43	2.6	107689	9	AL354681	AL354681	Human DNA	c 879	43	2.6	191446	2	AC024361	AC024361	Homo sapi
c 808	43	2.6	107895	9	AC078929	AC078929	Homo sapi	c 880	43	2.6	191446	2	AC105009	AC105009	Homo sapi
c 809	43	2.6	110000	2	AC067733_2	Continuation (3 of	c 881	43	2.6	191596	2	AC008658	AC008658	Homo sapi	
810	43	2.6	111551	9	AL360178	AL360178	Human DNA	c 882	43	2.6	193252	2	AC008658	AC008658	Homo sapi
811	43	2.6	111656	2	AC010449	AC010449	Homo sapi	c 883	43	2.6	193252	2	AC008658	AC008658	Homo sapi
812	43	2.6	114191	9	AL451007	AL451007	Human DNA	c 884	43	2.6	194167	9	AC093912	AC093912	Homo sapi
c 813	43	2.6	114517	9	AL158152	AL158152	Human DNA	c 885	43	2.6	194243	9	AC016821	AC016821	Homo sapi
c 814	43	2.6	123995	9	AC026248	AC026248	Homo sapi	c 886	43	2.6	194296	9	AL354864	AL354864	Human DNA
c 815	43	2.6	125780	9	CNS01DHW	AL137229	Human chr	887	43	2.6	196023	2	CNS01DHY	AL131594	Human chr
816	43	2.6	129192	9	HS204E5	298941	Human DNA s	c 888	43	2.6	196741	2	AC046185	AC046185	Homo sapi
817	43	2.6	129577	9	AC004659	AC004659	Homo sapi	c 889	43	2.6	197437	2	AF303736	AF303736	Homo sapi
818	43	2.6	132114	9	AC026772	AC026772	Homo sapi	c 890	43	2.6	201844	2	CNS05TCL	AL335102	Human chr
819	43	2.6	133769	9	AC010362	AC010362	Homo sapi	c 891	43	2.6	203912	9	AC008755	AC008755	Homo sapi
c 820	43	2.6	133769	9	AC010429	AC010429	Homo sapi	c 892	43	2.6	209574	2	AC040160	AC040160	Homo sapi
c 821	43	2.6	134957	9	HS9303L11	AL049759	Human DNA	c 893	43	2.6	211063	2	AC069254	AC069254	Homo sapi
c 822	43	2.6	137911	2	AC009769	AC009769	Homo sapi	894	43	2.6	218047	30	AC068624	AC068624	Homo sapi
823	43	2.6	140783	9	AC027373	AC027373	Homo sapi	c 895	43	2.6	221187	2	HS92N15	293097	Homo sapi
c 824	43	2.6	141737	2	AC124283	AC124283	Homo sapi	c 896	43	2.6	221187	2	AC023786	AC023786	Homo sapi
825	43	2.6	141899	2	AC079093	AC079093	Homo sapi	c 897	43	2.6	224048	2	AC008594	AC008594	Homo sapi
826	43	2.6	143444	9	AC007541	AC007541	Homo sapi	c 898	43	2.6	22516	2	AC008594	AC008594	Homo sapi
827	43	2.6	145414	2	AL135919	AL135919	Homo sapi	c 899	43	2.6	271144	9	HSXDPB	AL590763	Homo sapi
c 828	43	2.6	146652	2	AL445304	AL445304	Homo sapi	900	43	2.6	271203	2	AL772161	AL772161	Homo sapi
c 829	43	2.6	146903	2	AC024410	AC024410	Homo sapi	c 901	43	2.6	271814	2	AC027406	AC027406	Homo sapi
c 830	43	2.6	146903	2	AC024410	AC024410	Homo sapi	c 902	43	2.6	319138	2	AC091100	AC091100	Homo sapi
c 831	43	2.6	148131	9	AF248484	AF248484	Homo sapi	c 903	43	2.6	340000	9	HS21C007	AL163207	Homo sapi
c 832	43	2.6	149162	2	AC013758	AC013758	Homo sapi	904	42	2.5	313	9	HS9QT042	296779	H.sapiens t
c 833	43	2.6	149162	2	AC013758	AC013758	Homo sapi	905	42	2.5	420	9	HS16QT030	296330	H.sapiens t
c 834	43	2.6	151834	9	AP004195	AP004195	Homo sapi	c 906	42	2.5	424	11	HUMUT2349	LI8089	Human STS u
c 835	43	2.6	154065	2	AC084079	AC084079	Homo sapi	c 907	42	2.5	653	9	HSIGCL03B	X72827	H.sapiens (
c 836	43	2.6	154539	2	AC079187	AC079187	Homo sapi	c 908	42	2.5	1542	9	CLCN6HUM03	AF092924	Homo sapi
c 837	43	2.6	154708	2	AC024443	AC024443	Homo sapi	909	42	2.5	1887	9	AK092435	AK092435	Homo sapi
c 838	43	2.6	155742	9	CNS01DHW	AL137190	Human chr	910	42	2.5	5446	9	HSU78096	U78096	Human macro
c 839	43	2.6	156909	2	AB020863	AB020863	Homo sapi	c 911	42	2.5	6942	9	AB035185	AB035185	Homo sapi
c 840	43	2.6	157242	2	AC083821	AC083821	Homo sapi	c 912	42	2.5	7918	9	AB035184	AB035184	Homo sapi
c 841	43	2.6	157548	2	AC021972	AC021972	Homo sapi	913	42	2.5	16459	9	BSA297560	BSA297560	Homo sapi
c 842	43	2.6	157827	9	AL133387	AL133387	Human DNA	914	42	2.5	27613	9	AC002053	AC002053	Cosmid cl
c 843	43	2.6	157912	9	AL133387	AL133387	Human DNA	c 915	42	2.5	35877	9	AL392105	AL392105	Human DNA
c 844	43	2.6	158458	2	AC007395	AC007395	Homo sapi	c 916	42	2.5	35877	9	AL392105	AL392105	Human DNA
c 845	43	2.6	159335	2	AC046141	AC046141	Homo sapi	c 917	42	2.5	35878	9	AP000549	AP000549	Homo sapi
c 846	43	2.6	159535	2	AC046141	AC046141	Homo sapi	c 918	42	2.5	36008	9	AP001230	AP001230	Homo sapi
c 847	43	2.6	160004	9	AL139318	AL139318	Human DNA	919	42	2.5	38968	9	AC090098	AC090098	Homo sapi
c 848	43	2.6	163778	2	AC010868	AC010868	Homo sapi	c 920	42	2.5	38968	9	AC000075	AC000075	Homo sapi
c 849	43	2.6	164317	9	AL390029	AL390029	Human DNA	c 921	42	2.5	39000	9	HSN86D4	282250	Human DNA s
c 850	43	2.6	164824	2	AC073861	AC073861	Homo sapi	c 922	42	2.5	39372	9	HSN121F8	278421	Human DNA s
c 851	43	2.6	165139	2	AC008405	AC008405	Homo sapi	c 923	42	2.5	39854	9	AC004794	AC004794	Homo sapi
c 852	43	2.6	166264	2	AL390864	AL390864	Homo sapi	c 924	42	2.5	40981	9	HSB5E3	AL078611	Human DNA
c 853	43	2.6	168042	2	AC025931	AC025931	Homo sapi	c 925	42	2.5	41174	9	AC011547	AC011547	Homo sapi
c 854	43	2.6	168871	2	AC016117	AC016117	Homo sapi	c 926	42	2.5	41235	9	BSA512H23	AL121572	Human DNA
c 855	43	2.6	170470	9	AC092364	AC092364	Homo sapi	c 927	42	2.5	42265	9	AC007136	AC007136	Homo sapi
c 856	43	2.6	170862	2	AC087784	AC087784	Homo sapi	928	42	2.5	43420	2	AC002987	AC002987	Homo sapi
c 857	43	2.6	170896	2	AC011010	AC011010	Homo sapi	929	42	2.5	43336	9	AC010525	AC010525	Homo sapi
c 858	43	2.6	170940	2	AC025132	AC025132	Homo sapi	930	42	2.5	44504	9	AC022089	AC022089	Homo sapi
c 859	43	2.6	171075	2	AC026842	AC026842	Homo sapi	931	42	2.5	46324	2	AC093818	AC093818	Homo sapi
c 860	43	2.6	172008	9	AC079602	AC079602	Homo sapi	c 932	42	2.5	47226	2	AC105215	AC105215	Homo sapi
c 861	43	2.6	173109	2	AC091400	AC091400	Pan trogl	c 933	42	2.5	49773	9	AC099505	AC099505	Homo sapi
c 862	43	2.6	174818	9	AL390061	AL390061	Human DNA	c 934	42	2.5	51166	9	AC093701	AC093701	Homo sapi
c 863	43	2.6	174839	9	AL157395	AL157395	Human DNA	935	42	2.5	52314	9	AL627434	AL627434	Human DNA
c 864	43	2.6	176391	9	AL365400	AL365400	Human DNA	936	42	2.5	52884	2	AC027692	AC027692	Homo sapi
c 865	43	2.6	178340	2	AC022120	AC022120	Homo sapi	937	42	2.5	5966	2	AC124295	AC124295	Homo sapi
c 866	43	2.6	178499	2	AL360074	AL360074	Homo sapi	c 938	42	2.5	61904	2	AC091098	AC091098	Homo sapi
c 867	43	2.6	178635	2	AL591035	AL591035	Homo sapi	939	42	2.5	62271	9	AL390797	AL390797	Human DNA
c 868	43	2.6	180221	9	AC007277	AC007277	Homo sapi	940	42	2.5	63949	2	AC130686	AC130686	Homo sapi
							c 941	42	2.5	65237	2	AC105208	AC105208	Homo sapi	



942	c 943	42	2.5	65596	9	AC113432	AC113432	Homo sapi
943	c 944	42	2.5	65711	2	AC025892	AC025892	Homo sapi
944	c 945	42	2.5	65958	2	AC105413	AC105413	Homo sapi
945	c 946	42	2.5	66184	2	AC103841	AC103841	Homo sapi
946	c 947	42	2.5	66281	9	AL138710	AL138710	Human DNA
947	c 948	42	2.5	66586	2	AC090183	AC090183	Human DNA
948	c 949	42	2.5	66608	9	AL451063	AL451063	Human DNA
949	c 950	42	2.5	66836	2	AC016492	AC016492	Homo sapi
950	c 951	42	2.5	67550	2	AC016452	AC016452	Homo sapi
951	c 952	42	2.5	68004	2	AC104584	AC104584	Homo sapi
952	c 953	42	2.5	68676	2	AC026502	AC026502	Homo sapi
953	c 954	42	2.5	68697	2	AC131280	AC131280	Homo sapi
954	c 955	42	2.5	69523	2	AC124305	AC124305	Homo sapi
955	c 956	42	2.5	71365	2	AC130322	AC130322	Homo sapi
956	c 957	42	2.5	72097	2	AC130381	AC130381	Homo sapi
957	c 958	42	2.5	73656	2	AC016512	AC016512	Homo sapi
958	c 959	42	2.5	73656	2	AC016512	AC016512	Homo sapi
959	c 960	42	2.5	78000	9	AC115676	AC115676	Homo sapi
960	c 961	42	2.5	81419	9	AL133382	AL133382	Human DNA
961	c 962	42	2.5	82998	9	AP000967	AP000967	Homo sapi
962	c 963	42	2.5	83732	9	AC005183	AC005183	Homo sapi
963	c 964	42	2.5	84142	2	AP001110	AP001110	Homo sapi
964	c 965	42	2.5	84724	9	AC022084	AC022084	Homo sapi
965	c 966	42	2.5	85427	2	HS03470K3	HS03470K3	Homo sapi
966	c 967	42	2.5	86024	9	AL592103	AL592103	Human DNA
967	c 968	42	2.5	86516	2	AC022592	AC022592	Homo sapi
968	c 969	42	2.5	86719	9	AP000885	AP000885	Homo sapi
969	c 970	42	2.5	86743	9	AL359873	AL359873	Human DNA
970	c 971	42	2.5	86829	2	AC011382	AC011382	Homo sapi
971	c 972	42	2.5	86945	2	AC002490	AC002490	Homo sapi
972	c 973	42	2.5	87424	9	AC006077	AC006077	Homo sapi
973	c 974	42	2.5	87829	9	AC059333	AC059333	Homo sapi
974	c 975	42	2.5	87943	2	AC095064	AC095064	Homo sapi
975	c 976	42	2.5	88698	2	AC009008	AC009008	Homo sapi
976	c 977	42	2.5	90248	2	AC108366	AC108366	Pen trogl
977	c 978	42	2.5	90389	9	AC004008	AC004008	Homo sapi
978	c 979	42	2.5	91940	9	AL445674	AL445674	Human DNA
979	c 980	42	2.5	93011	9	AC090383	AC090383	Homo sapi
980	c 981	42	2.5	96299	9	AL133519	AL133519	Human DNA
981	c 982	42	2.5	96459	2	AC090233	AC090233	Homo sapi
982	c 983	42	2.5	96879	9	HS384F21	HS384F21	Human DNA
983	c 984	42	2.5	96978	9	AC024075	AC024075	Homo sapi
984	c 985	42	2.5	97015	9	AP002085	AP002085	Homo sapi
985	c 986	42	2.5	99489	9	AC005057	AC005057	Homo sapi
986	c 987	42	2.5	99489	9	AC008938	AC008938	Homo sapi
987	c 988	42	2.5	99593	9	HS7857M17	HS7857M17	Human DNA
988	c 989	42	2.5	100791	9	CNS01R18	CNS01R18	Human chr
989	c 990	42	2.5	100827	9	AC020909	AC020909	Homo sapi
990	c 991	42	2.5	101841	9	AL451185	AL451185	Human DNA
991	c 992	42	2.5	102591	9	AC002978	AC002978	Homo sapi
992	c 993	42	2.5	102818	9	AC003042	AC003042	Homo sapi
993	c 994	42	2.5	103009	9	AC087225	AC087225	Homo sapi
994	c 995	42	2.5	103606	9	AC005004	AC005004	Homo sapi
995	c 996	42	2.5	104892	9	AL589663	AL589663	Human DNA
996	c 997	42	2.5	104895	2	AC093288	AC093288	Homo sapi
997	c 998	42	2.5	104922	9	AF155238	AF155238	Homo sapi
998	c 999	42	2.5	104995	9	AC020942	AC020942	Homo sapi
999	c 1000	42	2.5	104995	9	AC026416	AC026416	Homo sapi
1000		42	2.5	105489	9	AC093894	AC093894	Homo sapi

RESULT 1  
AX394457

LOCUS  
AX394457

DEFINITION  
Sequence 2 from Patent WO0218638.

ACCESSION  
AX394457

VERSION  
AX394457.1

GI:21065595

KEYWORDS  
human.

SOURCE  
Homo sapiens

ORGANISM  
Homo sapiens

REFERENCE  
1  
Risinger, C., Andersson, M.K., Lewander, T. and Oliasson, E.  
Detection of cyp2d6 polymorphisms  
Patent: WO 0218638-A 2 07-MAR-2002;  
Gemini Genomics PLC (GB)

AUTHORS  
Risinger, C., Andersson, M.K., Lewander, T. and Oliasson, E.

TITLE  
Detection of cyp2d6 polymorphisms

JOURNAL  
Patent: WO 0218638-A 2 07-MAR-2002;  
Gemini Genomics PLC (GB)

FEATURES  
source  
1. .1680  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT  
413 a 379 c 539 g 342 t 7 others

ORIGIN

Query Match 100.0%; Score 1680; DB 6; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCAAGACCAGCCTGGACAACTTTGGAGAACCCSGGTCTCTACAAAAAATACAAAAAT 60  
Db 1 GAATTCAAGACCAGCCTGGACAACTTTGGAGAACCCSGGTCTCTACAAAAAATACAAAAAT 60  
QY 61 AGCTGGATTGGTGGCGGTGGCTCATGCCCTATATATCCAGCACTTTGGAGCCTGAGGTG 120  
Db 61 AGCTGGATTGGTGGCGGTGGCTCATGCCCTATATATCCAGCACTTTGGAGCCTGAGGTG 120  
QY 121 GGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCCTGGCCAACTGGTGAACCCCTATC 180  
Db 121 GGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCCTGGCCAACTGGTGAACCCCTATC 180  
QY 181 TCTACTGAAAAATAYAAAAAGCTAGAGCTGGTGGCACACACCTGTAATCCAGCTACTTAG 240  
Db 181 TCTACTGAAAAATAYAAAAAGCTAGAGCTGGTGGCACACACCTGTAATCCAGCTACTTAG 240  
QY 241 GAGGCTGAGCGAGGAGAAATGGCTTGAAGCCTAGAGCTGAAGGTTGTAGTGAGCCGAGATT 300  
Db 241 GAGGCTGAGCGAGGAGAAATGGCTTGAAGCCTAGAGCTGAAGGTTGTAGTGAGCCGAGATT 300  
QY 301 GCATCAATCCCAANTGAGGGAGCCACCAGCCTGGGCAACAAGAGGAATCTCCGCTC 360  
Db 301 GCATCAATGCACAATGGAGGGAGCCACCAGCCTGGGCAACAAGAGGAATCTCCGCTC 360  
QY 361 CAAAAAATAAAAAAATAAAAAAAGATTAGGCTGGTGGTGGCTGTAGTCCAGCTACTT 420  
Db 361 CAAAAAATAAAAAAATAAAAAAAGATTAGGCTGGTGGTGGCTGTAGTCCAGCTACTT 420  
QY 421 GGGAGCGAGGGGTCCACTTGATGTCGAGCTGAGCCATGATCTCCGCTGAC 480  
Db 421 GGGAGCGAGGGGTCCACTTGATGTCGAGCTGAGCCATGATCTCCGCTGAC 480  
QY 481 TCCGGCTGGGCAACAGCTGAGACCCCTCTCTAAAGAAAAAATAAAGCAACATATC 540  
Db 481 TCCGGCTGGGCAACAGCTGAGACCCCTCTCTAAAGAAAAAATAAAGCAACATATC 540  
QY 541 CTGAACAAAGGATCTCCATAACGTTCCACCAGATTCTTAATCGAAAAATCGAGGCCA 600  
Db 541 CTGAACAAAGGATCTCCATAACGTTCCACCAGATTCTTAATCGAAAAATCGAGGCCA 600  
QY 601 GAAAGCAGTGGAGGAGGACRACCCCTCAGCAGCCGGGAGGATGTTGTACAGGCTGGGG 660  
Db 601 GAAAGCAGTGGAGGAGGACRACCCCTCAGCAGCCGGGAGGATGTTGTACAGGCTGGGG 660  
QY 661 CAAGGCGCTTCGGGTACCAACTGGGAGCTCTGGGAACAGCCCTGTGCAAAACAAGAC 720  
Db 661 CAAGGCGCTTCGGGTACCAACTGGGAGCTCTGGGAACAGCCCTGTGCAAAACAAGAC 720  
QY 721 CATAGCCCGCCAGAGCCCGAGGAATGTGGCTGGGTGGGAGCAGCCTCTGACAGGACT 780  
Db 721 CATAGCCCGCCAGAGCCCGAGGAATGTGGCTGGGTGGGAGCAGCCTCTGACAGGACT 780  
QY 781 GGTCCCATCCAGAAACCTCCCGCATGGCTGGGAGTGGGGTACTTGGTGGCGGGTCTGT 840  
Db 781 GGTCCCATCCAGAAACCTCCCGCATGGCTGGGAGTGGGGTACTTGGTGGCGGGTCTGT 840



QY	841	ATGTGTGTGACTGGTGTGTGTGAGAGAGAATGTGTCYCTAAGTGTCAAGTGTGAGTCT	900
Db	841		
QY	901	GTGTATGTGTGAATATATGTTCTTGTGTGGTGATTTTCTGCRGTGTAAATCGTCCCTG	960
Db	901		
QY	961	CAAGTGTGAACAAGTGACAAAGTGTCTGGGAGTGGACAAGAGATCTGTGCACCATCAGGT	1020
Db	961		
QY	1021	GTGTGATACGCTCTGTGCATGTCAAGAGTGCAGAGTGAAGTGAAGGAGCAGGCCCATG	1080
Db	1021		
QY	1081	ATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGTAAAGTGCAGTGCACAGATAAGGGTG	1140
Db	1081		
QY	1141	CTGAAGGTCACTCTGGAGTGGGAGTGGGGTAGGGAAGGGCAAGGCCATGTTCTGGA	1200
Db	1141		
QY	1201	GGAGGGTTGTGACTACATTAGGTTGTATGAGCTTAGCTGGGAGTGGATGGCCRGCTCC	1260
Db	1201		
QY	1261	ACTGAACCCCTGGTTATCCAGAGGCTTTGCAGGCTTCAGGCTTCAGGAGTGGAGTGGGAGAG	1320
Db	1261		
QY	1321	GGGTGACTTCTCCGACAGGCCCTCCACCGGCTTACCCTGGGTAAAGGCCCTGGAGCAG	1380
Db	1321		
QY	1381	GAGCAGGGCAAGACCTCTGGAGCAGCCCATACCCGCCCTGGCCCTGACTCTGCCACTG	1440
Db	1381		
QY	1441	GCAGCAGTCAACACAGAGTTCACTCACAGCAGAGGCAAGGCCATCATCAGCTCC	1500
Db	1441		
QY	1501	CTTTATAAGGAAAGGTCACGCGCTCGGTGTGCTGAGAGTGTCTGCTGGTCTCTGTG	1560
Db	1501		
QY	1561	CCTGTGGGTGGGGTGCCAGGTGTGCCAGAGGCCATTTGGTGTGAGGCAGGTA	1620
Db	1561		
QY	1621	TGGGCTAGAGCAGTGGTCCCTGGCCCTGATAGTGGCCATCTTCTGCTCTGTGG	1680
Db	1621		
RESULT 2			
LOCUS	AX207224	1669 bp	DNA linear PAT 30-AUG-2001
DEFINITION	Sequence 1 from Patent WO0155432.		
ACCESSION	AX207224		
VERSION	AX207224.1 GI:15394976		
KEYWORDS	synthetic construct.		
SOURCE	synthetic construct.		
ORGANISM	artificial sequences.		
REFERENCE	1 (bases 1 to 1669)		
AUTHORS	Raimundo, S. and Zanger, U.		
TITLE	Polymorphisms in the human cyp2d6 gene promoter region and their use in diagnostic and therapeutic applications		
JOURNAL	Patent: WO 0155432-A 1 02-AUG-2001.		

FEATURES		Epidauros Biotechnologie AG (DE)										
source		Location/Qualifiers										
		1..1669										
		/organism="synthetic construct"										
		/db_xref="taxon:32630"										
		/note="artificial sequence"										
BASE COUNT	413 a	376 c	534 g	338 t	8 others							
ORIGIN												
Query Match		90.2%; Score 1516; DB 6; Length 1669;										
Best Local Similarity		99.8%; Pred. No. 0;										
Matches 1666; Conservative		0; Mismatches 3; Indels 0; Gaps 0;										
QY	1	GAATTC	AAAGC	ACCC	CTGG	ACAACTTGG	AAAGACCC	SGGTCTCT	CAAAAAAT	ACAAAATT	60	
Db	1	GAATTC	AAAGC	ACCC	CTGG	ACAACTTGG	AAAGACCC	SGGTCTCT	CAAAAAAT	ACAAAATT	60	
QY	61	AGCTGG	GATTTGG	GTGGCT	CATCGCTAT	AATCC	CAGCAC	ACTTTGG	GAGCCCT	GAGGTG	120	
Db	61	AGCTGG	GATTTGG	GTGGCT	CATCGCTAT	AATCC	CAGCAC	ACTTTGG	GAGCCCT	GAGGTG	120	
QY	121	GGTGG	ATCACCT	GAAGTC	AGGAGTTCA	AGACTAG	CCCTGG	CCCAACATG	GTG	AAACCCCTATC	180	
Db	121	GGTGG	ATCACCT	GAAGTC	AGGAGTTCA	AGACTAG	CCCTGG	CCCAACATG	GTG	AAACCCCTATC	180	
QY	181	TCTACT	GAAAAAT	YAAAAAG	CTAGAC	GTGGTG	CCACAC	ACCTGT	AAATCC	CAGCTACTTAG	240	
Db	181	TCTACT	GAAAAAT	YAAAAAG	CTAGAC	GTGGTG	CCACAC	ACCTGT	AAATCC	CAGCTACTTAG	240	
QY	241	GAGGCT	GAGG	CAGAG	AAATTCCTT	GAAGCC	TAGAG	GTGAAG	GTGTAG	TGAGCCGAGATT	300	
Db	241	GAGGCT	GAGG	CAGAG	AAATTCCTT	GAAGCC	TAGAG	GTGAAG	GTGTAG	TGAGCCGAGATT	300	
QY	301	GCATCAT	TGGACA	ATGGAG	GGGAGCC	ACCAGC	CTGGGCA	ACAACAG	GAGAAATCT	CCGCTCTC	360	
Db	301	GCATCAT	TGGACA	ATGGAG	GGGAGCC	ACCAGC	CTGGGCA	ACAACAG	GAGAAATCT	CCGCTCTC	360	
QY	361	CAAAAA	AAAAA	AGATTTAG	GCTGGGTGGT	GCCTGTAG	TCCC	AGCTACT	TTT	420		
Db	361	CAAAAA	AAAAA	AGATTTAG	GCTGGGTGGT	GCCTGTAG	TCCC	AGCTACT	TTT	420		
QY	421	GGAGG	CAGGGGT	TCCACTTGT	ATGTCG	AGACTGC	AGTGAG	CCATGAT	CTGCC	ACTTGCAC	480	
Db	421	GGAGG	CAGGGGT	TCCACTTGT	ATGTCG	AGACTGC	AGTGAG	CCATGAT	CTGCC	ACTTGCAC	480	
QY	481	TCCGG	CTGGCA	ACAGAG	TGAGAC	CTGTCTA	AAAGAAAA	AAAAAATA	AAAGCA	CACATATC	540	
Db	481	TCCGG	CTGGCA	ACAGAG	TGAGAC	CTGTCTA	AAAGAAAA	AAAAAATA	AAAGCA	CACATATC	540	
QY	541	CTGAAC	AAAGGAT	GCTCCAT	AAAGTTCCC	ACAGATTTCT	AATCAG	AAACATG	GAGGCCA	600		
Db	541	CTGAAC	AAAGGAT	GCTCCAT	AAAGTTCCC	ACAGATTTCT	AATCAG	AAACATG	GAGGCCA	600		
QY	601	GAAAG	CAGTGG	AGAGAC	RACCTC	AGGACGCC	CGGAGG	ATGTTGT	CA	AGGCTGGGG	660	
Db	601	GAAAG	CAGTGG	AGAGAC	RACCTC	AGGACGCC	CGGAGG	ATGTTGT	CA	AGGCTGGGG	660	
QY	661	CAAGG	CCCTTCC	GGCTAC	CAACTGG	GAGCTCT	GGGAAC	AGCCCTG	TTGCA	AAACAGAGC	720	
Db	661	CAAGG	CCCTTCC	GGCTAC	CAACTGG	GAGCTCT	GGGAAC	AGCCCTG	TTGCA	AAACAGAGC	720	
QY	721	CATAG	CCCGCC	AGAGCC	AGGAATGT	GGGTGG	GCTGG	GAGCAG	CCCTCT	GGACAGGAGT	780	
Db	721	CATAG	CCCGCC	AGAGCC	AGGAATGT	GGGTGG	GCTGG	GAGCAG	CCCTCT	GGACAGGAGT	780	
QY	781	GGTCC	CAATCC	AGGAAAC	CTCC	GGCATGG	CTGG	GAAGTGG	GCTACT	TGGTCC	CGGGTCTGT	840
Db	781	GGTCC	CAATCC	AGGAAAC	CTCC	GGCATGG	CTGG	GAAGTGG	GCTACT	TGGTCC	CGGGTCTGT	840
QY	841	ATGTG	TGTGTG	ACTGG	TGTGTGTG	AGAGAG	AAATGTGT	GCYCTA	AGTGTCA	AGTGTGAGTCT	900	
Db	841	ATGTG	TGTGTG	ACTGG	TGTGTGTG	AGAGAG	AAATGTGT	GCYCTA	AGTGTCA	AGTGTGAGTCT	900	



Db	997	CAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGCATGTCAAGAGTGC AAGG	1056
QY	1057	TGAAGTGAAGGGACCGCCATGATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGT	1116
Db	1057	TGAAGTGAAGGGACCGCCATGATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGT	1116
QY	1117	AAGTGCAGTGACAGATAAGGTGCTGAAGGTCACAGGTCACCTGAGTGGGAGCTGGGGTAGG	1176
Db	1117	AAGTGCAGTGACAGATAAGGGTGCTGAAGGTCACCTGAGTGGGAGCTGGGGTAGG	1176
QY	1177	GAAAGGGCAAGGCCATGTTCTGGAGGAGGGGTTGTGACTACATTAGGTCGTATCAGGCTTA	1236
Db	1177	GAAAGGGCAAGGCCATGTTCTGGAGGAGGGGTTGTGACTACATTAGGTCGTATCAGGCTTA	1236
QY	1237	GCTGGAGGTGGATGGCCRGCTCCACTGAAACCCCTGGTTATCCAGAAAGCTTTTCAGGC	1296
Db	1237	GCTGGAGGTGGATGGCCGGCTCCACTGAAACCCCTGGTTATCCAGAAAGCTTTTCAGGC	1296
QY	1297	TTCAGGAGCTTGAGTGGGAGAGGGGTCACCTTCTCCGACCAAGGCCCTCCACCGGCCT	1356
Db	1297	TTCAGGAGCTTGAGTGGGAGAGGGGTCACCTTCTCCGACCAAGGCCCTCCACCGGCCT	1356
QY	1357	ACCTCGGTAAGGGCCTGGAGCAGGAAGCAGGGGCAAGAACCTCTCGAGCAGGCCATACC	1416
Db	1357	ACCTCGGTAAGGGCCTGGAGCAGGAAGCAGGGGCAAGAACCTCTCGAGCAGGCCATACC	1416
QY	1417	CGCCCTGGCCTGACTGTGCACCTGGCAGCAGCAGTCAACACAGCAGGTTCACTCACAGCAG	1476
Db	1417	CGCCCTGGCCTGACTGTGCACCTGGCAGCAGCAGTCAACACAGCAGGTTCACTCACAGCAG	1476
QY	1477	AGGGCAAGGCCATCATCAGCTCCCTTTTATAAGGGAAGGTCAGCGCTCGGTGTGCTGA	1536
Db	1477	AGGGCAAGGCCATCATCAGCTCCCTTTTATAAGGGAAGGTCAGCGCTCGGTGTGCTGA	1536
QY	1537	GAGTGTCTGCCTGGTCTCTGTGCCTGGTGGGTGGGTGCCAGGTGTGCCAGAGGA	1596
Db	1537	GAGTGTCTGCCTGGTCTCTGTGCCTGGTGGGTGGGTGCCAGGTGTGCCAGAGGA	1596
QY	1597	GCCCATTTGGTACTGAGGCAGGTATGGGCTAGAAACACTGTGCCCCGTCGCCGTGATAG	1656
Db	1597	GCCCATTTGGTACTGAGGCAGGTATGGGCTAGAAACACTGTGCCCCGTCGCCGTGATAG	1656
QY	1657	TGCCCATCTTCCCTGCTCTGGTGG 1680	
Db	1657	TGCCCATCTTCCCTGCTCTGGTGG 1680	
RESULT 4			
HUMCYP2D6			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
COMMENT			
FEATURES			
Source			

gene	/map="22q13.1"	
	join(1532..1799,2503..2674,3225..3377,3466..3626,4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)	
mRNA	/gene="CYP2D6"	
	join(1532..1799,2503..2674,3225..3377,3466..3626,4060..4236,4427..4568,4776..4963,5418..5559,5658..5909)	
	/gene="CYP2D6"	
	/product="cytochrome P450 IID6"	
	/note="G00-132-127"	
exon	1532..1799	
	/partial	
	/gene="CYP2D6"	
	/note="cytochrome P450 IID6; G00-132-127"	
	/number=1	
CDS	join(1620..1799,2503..2674,3225..3377,3466..3626,4060..4236,4427..4568,4776..4963,5418..5559,5658..5836)	
	/gene="CYP2D6"	
	/codon_start=1	
	/product="cytochrome P450 IID6"	
	/protein_id="AAA53500.1"	
	/db_xref="GI:181304"	
	/db_xref="GB:G00-132-127"	
	/translation="MGLEALVLAIVAVLFLLLVDLMHRQRWAARYPPGPLEPLGLG	
	LLLVDFQNTPYCFDQLRRRFGDVFSQLAWTPVVLNGLAAVREALVTIGEDTADRP	
	PVITDILGFGRSQGVFLARYGPWREQRFRSVTLRNLGKKLSLEQWVTEAAACL	
	CAAFANHSVPLRLHIPALAGKVLRFKAFLTQDLLELTHRMTWMDPAQPRDLTEAPL	
	FIREVLNAVPLVLIIPALAGKVLRFKAFLTQDLLELTHRMTWMDPAQPRDLTEAPL	
	AMEKAKNPESSEFNENLRIVVADLFSAGMVTSTTLAGMLLMLTHPDVORRVOQE	
	IDVIGQVRPEMGDOAHMPTTAVTHEVQRFGDIVPLGVTHTMSRDIEVQGFRIPKG	
	TLTLNLSVLDKAEAWKPFRRFHPHFHLDAGQHFVKPEAFLEFSSAGRACCLGEPLAR	
	MELFLFTSLLOHFSFVSPTGQPRPSHHGVFAFLVSPSYELCAVPR"	
	1800..2502	
intron	/gene="CYP2D6"	
	/note="G00-132-127"	
	/number=1	
exon	2503..3626	
	/gene="CYP2D6"	
	/note="G00-132-127"	
intron	/number=2	
	2675..3224	
	/gene="CYP2D6"	
	/note="G00-132-127; does not fit consensus"	
exon	/number=2	
	3225..3377	
	/gene="CYP2D6"	
	/note="G00-132-127"	
	/number=3	
intron	3378..3465	
	/gene="CYP2D6"	
	/note="G00-132-127"	
exon	/number=3	
	3466..3626	
	/gene="CYP2D6"	
	/note="G00-132-127"	
intron	/number=4	
	3627..4059	
	/gene="CYP2D6"	
	/note="G00-132-127"	
	/number=4	
exon	4060..4236	
	/gene="CYP2D6"	
	/note="G00-132-127"	
intron	/number=5	
	4237..4426	
	/gene="CYP2D6"	
	/note="G00-132-127"	
exon	/number=5	
	4427..4568	
	/gene="CYP2D6"	
	/note="G00-132-127"	
intron	/number=6	
	4569..4775	
	/gene="CYP2D6"	



DEFINITION Human CYP2D7BP pseudogene for cytochrome P450 2D6.  
ACCESSION X58468  
VERSION X58468.1 GI:30337  
KEYWORDS CYP2D7BP gene; Cytochrome P450; cytochrome P450 2D6; pseudogene.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
Helm,M.H.  
Direct Submission  
Submitted (25-MAR-1991) M.H. Helm, Dept of Pharmacology, Biocentre  
University of Basel, Klingelbergstr 70, 4056 Basel, SWITZERLAND  
2 (bases 1 to 13677)  
Helm,M.H. and Meyer,U.A.  
Evolution of a highly polymorphic human cytochrome P450 gene  
cluster: CYP2D6  
Genomics 14 (1), 49-58 (1992)  
MEDLINE 93052308  
PubMed 1358797  
COMMENT See X58467, and Am. J. Hum. Genet. 47:994-1001(1990).  
FEATURES  
source  
1..13677  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/clone="45"  
/clone\_lib="EMBL4"  
/dev\_stage="adult"  
1534..5868  
/gene="CYP2D7BP"  
Join(1534..1801,2504..2675,3203..3355,3444..3605,  
4031..4207,4400..4540,4735..4922,5377..5518,5617..5868)  
/gene="CYP2D7BP"  
1534..1801  
/gene="CYP2D7BP"  
/number=1  
Join(1622..1801,2504..2675,3203..3355,3444..3605,  
4031..4207,4400..4540,4735..4922,5377..5518,5617..5795)  
/gene="CYP2D7BP"  
/pseudo  
/codon\_start=1  
/db\_xref="PID:e34293"  
1802..2503  
/gene="CYP2D7BP"  
/number=1  
2504..2675  
/gene="CYP2D7BP"  
/number=2  
2676..3202  
/gene="CYP2D7BP"  
/number=2  
3203..3355  
/gene="CYP2D7BP"  
/number=3  
3356..3443  
/gene="CYP2D7BP"  
/number=3  
3444..3605  
/gene="CYP2D7BP"  
/number=4  
3606..4030  
/gene="CYP2D7BP"  
/number=4  
4031..4207  
/gene="CYP2D7BP"  
/number=5  
4208..4399  
/gene="CYP2D7BP"  
/number=5  
4400..4540  
/gene="CYP2D7BP"  
/number=6

intron 4541..4734  
/gene="CYP2D7BP"  
/number=6  
exon 4735..4922  
/gene="CYP2D7BP"  
/number=7  
intron 4923..5376  
/gene="CYP2D7BP"  
/number=7  
exon 5377..5518  
/gene="CYP2D7BP"  
/number=8  
intron 5519..5616  
/gene="CYP2D7BP"  
/number=8  
exon 5617..5868  
/gene="CYP2D7BP"  
/number=9  
repeat\_region 8267..8306  
/note="ACCTTCCCC"  
BASE COUNT 3066 a 3775 c 4107 g 2729 t  
ORIGIN  
Query Match 68.9%; Score 1158; DB 9; Length 13677;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 6; Indels 2; Gaps 1;  
QY 37 GGTCCTACAAAAATACAAAATAGCTGGGATTGGGTGGCTCATGCGCTCATGCCTATATATC 96  
Db 37 GGTCCTACAAAAATACAAAATAGCTGGGATTGGGTGGCTCATGCGCTATATATC 96  
QY 97 CCAGCAGCTTTGGGAGCCTGAGGTGGGTGGATCACCTGAAGTCAAGGAGTTCAAGACTAGCC 156  
Db 97 CCAGCAGCTTTGGGAGCCTGAGGTGGGTGGATCACCTGAAGTCAAGGAGTTCAAGACTAGCC 156  
QY 157 TGCCCAACATGGTGAACCCCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGGCAC 216  
Db 157 TGCCCAACATGGTGAACCCCTATCTCTACTGAAATATAAAAAGCTAGACGTGGTGGCAC 216  
QY 217 ACACCTGTAATCCAGCTACTTAGGAGGCTGAGCGAGGAGAAATGCTTGAAGCCTAGAGG 276  
Db 217 ACACCTGTAATCCAGCTACTTAGGAGGCTGAGCGAGGAGAAATGCTTGAAGCCTAGAGG 276  
QY 277 TGAAGTTGTAGTGAGCCGAGATTGCATCATTTGCACAAATGGAGGGAGCCACCAGCCTGG 336  
Db 277 TGAAGTTGTAGTGAGCCGAGATTGCATCATTTGCACAAATGGAGGGAGCCACCAGCCTGG 336  
QY 337 GCAACAAGAGGAAATCTCCGCTCC--AAAAAAGGATTTAGGCTG 394  
Db 337 GCAACAAGAGGAAATCTCCGCTCCAAAAAAGGATTTAGGCTG 396  
QY 395 GGTGGTGCCTGTAGTCCCAGCTACTTGGGAGGAGGCGGTCCACTTCATGTCGAGACTGC 454  
Db 397 GGTGGTGCCTGTAGTCCCAGCTACTTGGGAGGAGGCGGTCCACTTCATGTCGAGACTGC 456  
QY 455 AGTGAGCCATGATCTCTGCCACTGCCTCCGGCCTGGGCAACAGAGTGAGACCTGTCTAA 514  
Db 457 AGTGAGCCATGATCTCTGCCACTGCCTCCGGCCTGGGCAACAGAGTGAGACCTGTCTAA 516  
QY 515 AGAAAAAATAAAGCAACATATCTTGAACAAAGGATCCTCCATAACGTTCCCAACAG 574  
Db 517 AGAAAAAATAAAGCAACATATCTTGAACAAAGGATCCTCCATAACGTTCCCAACAG 576  
QY 575 ATTTCCTAATCAGAAACATGGAGCCAGAAAGCAGTGGAGGAGGACACCTCCTCAGGAGCC 634  
Db 577 ATTTCCTAATCAGAAACATGGAGCCAGAAAGCAGTGGAGGAGGACACCTCCTCAGGAGCC 636  
QY 635 CGGAGGATGTTGTACAGGCTGGGCAAGGCGCTTCCGGCTTACCAACTGGGAGCTTGG 694  
Db 637 CGGAGGATGTTGTACAGGCTGGGCAAGGCGCTTCCGGCTTACCAACTGGGAGCTTGG 696  
QY 695 GAACAGCCCTGTTGCAACAAAGAGCCATAGCCCGGCCAGAGCCCAAGGAATGTGGGCTGG 754  
|||||

Db	697	GAACAGCCCTGTTGCAACAAGAACCATAGCCCGCCAGACGCCAGGAAATGTGGCTGG	756
QY	755	GCTGGAGACAGCCTCTGGACAGGAGTGGTCCATCAGAGAAACCTCCGGCATGGCTGGGA	814
Db	757	GCTGGAGACAGCCTCTGGACAGGAGTGGTCCATCAGAGAAACCTCCGGCATGGCTGGGA	816
QY	815	AGTGGGGTACTTGGTGGCCGGGTCCTGATGCTGTGTGACTGCTGTGTGAGAGAGAAATG	874
Db	817	AGTGGGGTACTTGGTGGCCGGGTCCTGATGCTGTGTGACTGCTGTGTGAGAGAGAAATG	876
QY	875	TGTGCTTAAGTGTTCAGTGTGAGTCTGTGTATGTGTGAATATGTCTTTGTGGGTGAT	934
Db	877	TGTGCCCTAAGTGTTCAGTGTGAGTCTGTGTATGTGTGAATATGTCTTTGTGGGTGAT	936
QY	935	TTTCTGCRGTGTAATCTGTGCCCTGCAAGTGTGAACAAGTGGACAAAGTGTCTGGAGATG	994
Db	937	TTTCTGCRGTGTAATCTGTGCCCTGCAAGTGTGAACAAGTGGACAAAGTGTCTGGAGATG	996
QY	995	GACAAGAGATCTGTCCACCATCAGTGTGTGCATAGCGTCTGTGCATCTCAAGAGTGCAA	1054
Db	997	GACAAGAGATCTGTCCACCATCAGTGTGTGCATAGCGTCTGTGCATCTCAAGAGTGCAA	1056
QY	1055	GGTGAAGTGAAGGGACCAAGGCCATGATGCCACTCATCATCAGGAGCTCTTAAGGCCCCAG	1114
Db	1057	GGTGAAGTGAAGGGACCAAGGCCATGATGCCACTCATCATCAGGAGCTCTTAAGGCCCCAG	1116
QY	1115	GTAAGTGGCAGTGACAGATAAGGTGCTGAAGTGCACCTCTGGAGTGGCAGGTGGGGGTA	1174
Db	1117	GTAAGTGGCAGTGACAGATAAGGTGCTGAAGTGCACCTCTGGAGTGGCAGGTGGGGGTA	1176
QY	1175	GGGAAAGGGCAAGGCCATGTTCTGGAGGAGGGGTGTGACTACATTAGGGTGTATGAGCC	1234
Db	1177	GGGAAAGGGCAAGGCCATGTTCTGGAGGAGGGGTGTGACTACATTAGGGTGTATGAGCC	1236
QY	1235	TAGCTGGAGGTGGAATGGCCRGTCCTGAACCCCTGGTTATCCAGAAAGCTTTTGCAAG	1294
Db	1237	TAGCTGGAGGTGGAATGGCCRGTCCTGAACCCCTGGTTATCCAGAAAGCTTTTGCAAG	1296
QY	1295	GCTTCAGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACCAAGGCCCTCCACCGGC	1354
Db	1297	GCTTCAGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACCAAGGCCCTCCACCGGC	1356
QY	1355	CTACCTCGGTGAAGGCCCTGGACAGGAAGCAGGGGCAAGAACCTCTGGAGCAGCCCAT	1414
Db	1357	CTACCTCGGTGAAGGCCCTGGACAGGAAGCAGGGGCAAGAACCTCTGGAGCAGCCCAT	1416
QY	1415	CCCGCCCTGGCCTGACTCTGCCACTGGCAGCACAGTCAACACACAGGTTCCACTCACAGC	1474
Db	1417	CCCGCCCTGGCCTGACTCTGCCACTGGCAGCACAGTCAACACACAGGTTCCACTCACAGC	1476
QY	1475	AGAGGGCAAAAGGCCATCATCAGTCCCTTTTATAAGGGAAGGTCACCGCTCTGGTGTGCT	1534
Db	1477	AGAGGGCAAAAGGCCATCATCAGTCCCTTTTATAAGGGAAGGTCACCGCTCTGGTGTGCT	1536
QY	1535	GAGAGTGTCCCTGGCTGGTCCCTCTGTGCTGTGGGTGGGGGTGCCAGGTGTGTCCAGAG	1594
Db	1537	GAGAGTGTCCCTGGCTGGTCCCTCTGTGCTGTGGGTGGGGGTGCCAGGTGTGTCCAGAG	1596
QY	1595	GAGCCCAATTTGGTGTAGTCAGGCAGGTATGGGGCTAGAAGCACTGGTGGCCCTGGCCGTGAT	1654
Db	1597	GAGCCCAATTTGGTGTAGTCAGGCAGGTATGGGGCTAGAAGCACTGGTGGCCCTGGCCGTGAT	1656
QY	1655	AGTGGCCATCTTCCTGCTCTGTGGTGG	1680
Db	1657	AGTGGCCATCTTCCTGCTCTGTGGTGG	1682
RESULT 6			
HUMCYP2D6			
LOCUS	HUMCYP2D6	5503 bp	DNA linear PRI 27-APR-1993
DEFINITION	Human debrisoquine 4-hydroxylase mutant allele (CYP2D6-MA1) gene, complete cds.		
ACCESSION	M33189		

M33189.1 GI:181305  
debrisoquine 4-hydroxylase.  
Human individual MAGA DNA.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 5503)  
Gonzalez,F.J.  
Unpublished (1990)  
Draft entry and computer-readable sequence for [1] kindly submitted  
by F.Gonzalez, 23-MAR-1990, for release after publication.  
Author address: F.Gonzalez  
National Cancer Institute  
Bldg. 37 Rm. 3E-24  
National Institute of Health  
Bethesda, Md 20892.  
FEATURES  
Location/Qualifiers  
1..5503  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
TATA\_signal  
prim\_transcript 726..5103  
CDS  
1join(814..993,1696..1877,2419..2571,2661..2820,3254..3430).  
3621..3762,3970..4157,4612..4753,4852..5030)  
/note="debrisoquine 4-hydroxylase"  
/codon\_start=1  
/protein\_id="AAA35737.1"  
/db\_xref="GI:181306"  
/translation="MGLFALPLAVIVAIFLLVLDLMHRRQRAARYSPGGLPLGIG  
NLLHVDFTQPCYCDQLRRRGDVFSLQAWTPVVLNGLAAVREAMVRGETADRRP  
PVPITQILFGFRSGQKRGVPGALWARVAEALLRLHLAQLGPGQVGAAGVDRGG  
RLPLCRQLQRPVFRPNGLLDKAVSNVIASTLCGRFFEYDDPRFLDLDAQEGLKE  
ESGLELRVAPVLLHIPALAGVLRFQKAFLTQLDELTEHRMTWDPAPQPRDLTE  
AFLAEMERAKGPNSSNDENLRVADLFSAGAVTTSTTLAWGLLIMLHPDQVRV  
QOEIDDVQVRREMGDQAHMPTTAVIHEVQREGDIVPLGVTMTSRDIEVQGFRI  
PKGTTLITNLSSVLKDAVWKPRFPHDFLDAQGHFVKPEAFLPESARRACLGEF  
LARLELFLFTSLQHFSEFVSPTQPRPSHHGVAFVLVTPSELCAVPR"  
<814..993  
/note="debrisoquine 4-hydroxylase"  
/number=1  
994..1695  
/note="debrisoquine 4-hydroxylase intron A"  
1696..1877  
/number=2  
1878..2418  
/note="debrisoquine 4-hydroxylase intron B"  
2419..2571  
/number=3  
2572..2660  
/note="debrisoquine 4-hydroxylase intron C"  
2661..2820  
/number=4  
2821..3253  
/note="debrisoquine 4-hydroxylase intron D"  
3254..3430  
/number=5  
3431..3620  
/note="debrisoquine 4-hydroxylase intron E"  
3621..3762  
/number=6  
3763..3969  
/note="debrisoquine 4-hydroxylase intron F"  
3970..4157  
/number=7  
4158..4611  
/note="debrisoquine 4-hydroxylase intron G"  
4612..4753  
/number=8  
4754..4851  
/note="debrisoquine 4-hydroxylase intron H"  
4852..>5030  
/note="debrisoquine 4-hydroxylase"

BASE COUNT		1066 a	1537 c	1851 g	1049 t		
ORIGIN	Chromosome 22.						
Query Match		42.9%; Score 721; DB 9; Length 5503;					
Best Local Similarity		99.7%; Pred. No. 0;					
Matches 871; Conservative		0; Mismatches 3; Indels 0; Gaps 0;					
QY	807	GGCTGGGAAGTGGGGTACTTGGTGGCCGGGTCTGTATGTGTGTGACTGGTGTGTGAG	866				
Db	1	GGCTGGGAAGTGGGGTACTTGGTGGCCGGGTCTGTATGTGTGTGACTGGTGTGTGAG	60				
QY	867	AGAAATGTGTCYCAAGTGTCAAGTGTGAGTCTGTATGTGTGAATATTGCTTTTGTG	926				
Db	61	AGAAATGTGTGCCCTAAAGTGTGAGTCTGTATGTGTGAATATTGCTTTTGTG	120				
QY	927	TGGGTGATTTTCTGCTGTATCTGTAATCTGTCTCCCTGCAAGTGTGAACAAGTGGACAAGTGT	986				
Db	121	TGGGTGATTTTCTGCTGTGTATCTGTCTCCCTGCAAGTGTGAACAAGTGGACAAGTGT	180				
QY	987	TGGAGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGATGTCAA	1046				
Db	181	TGGAGTGGACAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGATGTCAA	240				
QY	1047	GAGTGAAGGTGAAGTGAAGGACCAAGCCCATGATGCCACTCATCAGGAGCTCTAA	1106				
Db	241	GAGTGAAGGTGAAGTGAAGGACCAAGCCCATGATGCCACTCATCAGGAGCTCTAA	300				
QY	1107	GGCCCCAGGTGAAGTGCACGTACAGATAGGGGTGCTGAAGGTGCACCTCGAGTGGCAGG	1166				
Db	301	GGCCCCAGGTGAAGTGCACGTACAGATAGGGGTGCTGAAGGTGCACCTCGAGTGGCAGG	360				
QY	1167	TGGGGGTAGGAAAGGCAAGGCCATGTCTCTGGAGGAGGGTGTGTGACATACATTAGGGTG	1226				
Db	361	TGGGGGTAGGAAAGGCAAGGCCATGTCTCTGGAGGAGGGTGTGTGACATACATTAGGGTG	420				
QY	1227	TATGACCCCTAGCTGGAGGTGGATGCCRGCTCCATGAAACCCCTGGTTATCCCAAGG	1286				
Db	421	TATGACCCCTAGCTGGAGGTGGATGCCRGCTCCATGAAACCCCTGGTTATCCCAAGG	480				
QY	1287	CTTTGCAGGCTTCAGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACGAGCCCT	1346				
Db	481	CTTTGCAGGCTTCAGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACGAGCCCT	540				
QY	1347	CCACCGGCTACCCCTGGGTGAAGGCTTGGAGCAGGAAGCAGGGGCAAGAACCTCTGGAGC	1406				
Db	541	CCACCGGCTACCCCTGGGTGAAGGCTTGGAGCAGGAAGCAGGGGCAAGAACCTCTGGAGC	600				
QY	1407	AGCCCATACCCGCTTGGCTGACTCTGCCACTGGCAGCACAGTCAACACAGCAGGTTC	1466				
Db	601	AGCCCATACCCGCTTGGCTGACTCTGCCACTGGCAGCACAGTCAACACAGCAGGTTC	660				
QY	1467	CTCACAGCAGAGGGCAAGGCCATCATCAGTCCCTTTTATAAGGGAAGGTCACGCGCTC	1526				
Db	661	CTCACAGCAGAGGGCAAGGCCATCATCAGTCCCTTTTATAAGGGAAGGTCACGCGCTC	720				
QY	1527	GGTGTCTCAGAGTGTCTCCCTGCTGTGTCTGTGTGTGGGTGGGGTGGCCAGGTGT	1586				
Db	721	GGTGTCTCAGAGTGTCTCCCTGCTGTGTCTGTGTGTGGGTGGGGTGGCCAGGTGT	780				
QY	1587	GTCACAGAGAGCCCATTTGGTAGTAGGAGGATGGGGCTAGAACACTGTGGCCCTG	1646				
Db	781	GTCACAGAGAGCCCATTTGGTAGTAGGAGGATGGGGCTAGAACACTGTGGCCCTG	840				
QY	1647	GCGGTGATAGTGGCCATCTTCCTGCTCTCTGGTGG	1680				
Db	841	GCGGTGATAGTGGCCATCTTCCTGCTCTCTGGTGG	874				

RESULT 7  
HSCYP2D7A HSCYP2D7A 13278 bp DNA linear PRI 21-OCT-1992  
LOCUS Human CYP2D7AP pseudogene for cytochrome p450 2D6.  
DEFINITION

ACCESSION	X58467	
VERSION	X58467.1	GI:30336
KEYWORDS	CYP2D7AP gene; Cytochrome P450; cytochrome P450 2D6; pseudogene.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	
Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;		
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.		
1 (bases 1 to 13278)		
Heim,M.H.		
Direct Submission		
Submitted (25-MAR-1991) M.H. Heim, Dept of Pharmacology, Biocentre		
University of Basel, Klingelbergstr 70, 4056 Basel, SWITZERLAND		
2 (bases 1 to 13278)		
Heim,M.H. and Meyer,U.A.		
Evolution of a highly polymorphic human cytochrome P450 gene		
cluster: CYP2D6		
Genomics 14 (1), 49-58 (1992)		
93052308		
1358797		
See X58468, and Am. J. Hum. Genet. 47:994-1001(1990).		
Location/Qualifiers		
1. .13278		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/chromosome="22"		
/clone_lib="EMBL4"		
/dev_stage="adult"		
1154. .5489		
/gene="CYP2D7AP"		
Join(1154. .1424,2125. .2296,2823. .2975,3064. .3225,		
3651. .3827,4020. .4161,4356. .4542,4998. .5139,5238. .5489)		
/gene="CYP2D7AP"		
1154. .1424		
/gene="CYP2D7AP"		
/number=1		
Join(1242. .1424,2125. .2296,2823. .2975,3064. .3225,		
3651. .3827,4020. .4161,4356. .4542,4998. .5139,5238. .5416)		
/gene="CYP2D7AP"		
/pseudo		
/codon_start=1		
/db_xref="PID:e34292"		
1425. .2124		
/gene="CYP2D7AP"		
/number=1		
2125. .2296		
/gene="CYP2D7AP"		
/number=2		
2297. .2822		
/gene="CYP2D7AP"		
/number=2		
2823. .2975		
/gene="CYP2D7AP"		
/number=3		
2976. .3063		
/gene="CYP2D7AP"		
/number=3		
3064. .3225		
/gene="CYP2D7AP"		
/number=4		
3226. .3650		
/gene="CYP2D7AP"		
/number=4		
3651. .3827		
/gene="CYP2D7AP"		
/number=5		
3828. .4019		
/gene="CYP2D7AP"		
/number=5		
4020. .4161		
/gene="CYP2D7AP"		
/number=6		
4162. .4355		

gene  
mRNA  
exon  
CDS  
intron  
exon  
intron  
exon  
intron  
exon  
intron  
exon  
intron  
exon  
intron  
exon  
intron



/gene="CYP2D7AP"  
/number=6  
4356..4542  
/gene="CYP2D7AP"  
/number=7  
4543..4997  
/gene="CYP2D7AP"  
/number=7  
4998..5139  
/gene="CYP2D7AP"  
/number=8  
5140..5237  
/gene="CYP2D7AP"  
/number=8  
5238..5489  
/gene="CYP2D7AP"  
/number=9  
7887..7908  
repeat\_region  
/note="ACCCCTCCCC repeat"  
BASE COUNT 2902 a 3664 c 3968 g 2744 t  
ORIGIN

Query Match 17.1%; Score 287; DB 9; Length 13278;  
Best Local Similarity 99.7%; Pred. No. 9.4e-138;  
Matches 337; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 884 AGTGTCAAGTGTGAGTCTGTGTATGTGTAATATGTTTGTGGTGAATTTCTGCGT 943  
|||||  
Db 506 AGTGTCAAGTGTGAGTCTGTGTATGTGTAATATGTTTGTGGTGAATTTCTGCGT 555  
QY 944 GTGTATTCGTGTCCTGCAAGTGTGAACAAAGTGTCTGGAGTGGACAAGAGA 1003  
|||||  
Db 566 GTGTATTCGTGTCCTGCAAGTGTGAACAAAGTGTCTGGAGTGGACAAGAGA 625  
QY 1004 TCTGTGCACCATCAGGTGTGTCATAGCGTCTGTGATGTCACAGAGTGCAGGTGAAGTG 1063  
|||||  
Db 626 TCTGTGCACCATCAGGTGTGTCATAGCGTCTGTGATGTCACAGAGTGCAGGTGAAGTG 685  
QY 1064 AAGGGACAGGCCCATGATCCCACTCATCATCAGGAGCTTAAGGCCCAAGTAAAGTGC 1123  
|||||  
Db 686 AAGGGACAGGCCCATGATCCCACTCATCATCAGGAGCTTAAGGCCCAAGTAAAGTGC 745  
QY 1124 AGTGACAGATAAGGGTCTGAAAGTCACTCTGGAGTGGCAGGTGGGGTAGGAAAGG 1183  
|||||  
Db 746 AGTGACAGATAAGGGTCTGAAAGTCACTCTGGAGTGGCAGGTGGGGTAGGAAAGG 805  
QY 1184 CAAGGCCATGTTCTGGAGGAGGGTTGTGACTACATTA 1221  
|||||  
Db 806 CAAGGCCATGTTCTGGAGGAGGGTTGTGACTACATTA 843

RESULT 8  
HUMCYP8P HUMCYP8P 17060 bp DNA linear PRI 09-NOV-1994  
LOCUS Human debrisoquine 4-hydroxylase (CYP2D8P) and (CYP2D7) pseudogenes  
DEFINITION complete sequences.  
ACCESSION M33387  
VERSION 1 GI:181320  
KEYWORDS debrisoquine 4-hydroxylase.  
SOURCE Human DNA, clones lambda-2D-A and lambda-2D-B.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 17060)  
AUTHORS Kimura S., Umeno M., Skoda R.C., Meyer, U.A. and Gonzalez, F.J.  
TITLE The human debrisoquine 4-hydroxylase (CYP2D) locus: sequence and  
identification of the polymorphic CYP2D6 gene, a related gene, and  
a pseudogene  
JOURNAL Am. J. Hum. Genet. 45 (6), 889-904 (1989)  
MEDLINE 90072069  
PUBMED 2574001  
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted  
by S.Kimura, 29-MAR-1990.

FEATURES  
source Location/Qualifiers  
1..17060  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
TATA\_signal  
prim\_transcript 1276..1282  
1304..6570  
/note="CYP2D8P mRNA and introns"  
gene join(1392..1568,3189..3360,3907..4059,4148..4310,  
4758..4934,5121..5262,5467..5651,6101..6242,6339..6516)  
/gene="CYP2D8P"  
CDS join(1392..1568,3189..3360,3907..4059,4148..4310,  
4758..4934,5121..5262,5467..5651,6101..6242,6339..6516)  
/gene="CYP2D8P"  
/pseudo  
/codon\_start=1  
/product="debrisoquine 4-hydroxylase"  
exon <1392..1568  
/gene="CYP2D8P"  
exon 3189..3360  
/number=1  
/gene="CYP2D8P"  
exon 3907..4059  
/number=2  
/gene="CYP2D8P"  
exon 4148..4310  
/number=3  
/gene="CYP2D8P"  
exon 4758..4934  
/number=4  
/pseudo  
/gene="CYP2D8P"  
exon 5121..5262  
/number=5  
/pseudo  
exon 5467..5651  
/gene="CYP2D8P"  
exon 6101..6242  
/number=6  
/gene="CYP2D8P"  
exon 6339..>6516  
/number=7  
/gene="CYP2D8P"  
exon 6516..6570  
/number=8  
TATA\_signal 11209..11215  
prim\_transcript 11236..11571  
gene join(11324..11504,12206..12377,12906..13058,13147..13307,  
13733..13909,14102..14243,14438..14625,15080..15221,  
15320..15496)  
/gene="CYP2D7"  
CDS join(11324..11504,12206..12377,12906..13058,13147..13307,  
13733..13909,14102..14243,14438..14625,15080..15221,  
15320..15496)  
/gene="CYP2D7"  
exon <11324..11504  
/number=1  
exon 12206..12377  
/gene="CYP2D7"  
exon 12906..13058  
/pseudo  
/gene="CYP2D7"  
exon <11324..11504  
/number=2  
/product="debrisoquine 4-hydroxylase"  
exon <11324..11504  
/number=1  
exon 12206..12377  
/gene="CYP2D7"  
exon 12906..13058  
/pseudo  
/gene="CYP2D7"



exon /number=3  
/pseudo  
13147..13307  
/gene="CYP2D7"  
/number=4  
/pseudo  
13733..13909  
/gene="CYP2D7"  
/number=5  
/pseudo  
14102..14243  
/gene="CYP2D7"  
/number=6  
/pseudo  
14438..14625  
/gene="CYP2D7"  
/number=7  
/pseudo  
15080..15221  
/gene="CYP2D7"  
/number=8  
/pseudo  
15320..>15496  
/gene="CYP2D7"  
/number=9

BASE COUNT 3516 a 4595 c 5034 g 3915 t  
ORIGIN

Query Match 16.0%; Score 269; DB 9; Length 17060;  
Best Local Similarity 99.5%; Pred. No. 1.9e-128;  
Matches 369; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 884 AGTGTCAAGTGTGAGTGTGATGTGTAATATGTTTGTGTGGTGATTTTCTGCRT 943  
|||||  
Db 10585 AGTGTCAAGTGTGAGTGTGATGTGTAATATGTTTGTGTGGTGATTTTCTGCAT 10644  
|||||  
QY 944 GTGTAATCGTGTCCCTGCAAGTGTGAACAAGTGACAAAGTGTCTGGAGTGGACAAGAGA 1003  
|||||  
Db 10645 GTGTAATCGTGTCCCTGCAAGTGTGAACAAGTGTCTGGAGTGGACAAGAGA 10704  
|||||  
QY 1004 TCTGTGCACCATCAGGTGTGTGATGCTGTGTCATGCTGTGTCATGCAAGAGTGCAGGTGAAGTG 1063  
|||||  
Db 10705 TCTGTGCACCATCAGGTGTGTGATGCTGTGTCATGCTGTGTCATGCAAGAGTGCAGGTGAAGTG 10764  
|||||  
QY 1064 AAGGACAGGCCCATGATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGTAAGTGCC 1123  
|||||  
Db 10765 AAGGACAGGCCCATGATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGTAAGTGCC 10824  
|||||  
QY 1124 AGTCACAGATAAGGGTGCTCAAGTGTCACTCTGGAGTGGGCGAGGTGGGGTAGGAAAGGG 1183  
|||||  
Db 10825 AGTCACAGATAAGGGTGCTCAAGTGTCACTCTGGAGTGGGCGAGGTGGGGTAGGAAAGGG 10884  
|||||  
QY 1184 CAAGGCCATGCTTCTGGAGGAGGGTGTGTGACTACTATAGGCTGTATGAGCCTAGCTGGGA 1243  
|||||  
Db 10885 CAAGTGTATGCTTCTGGAGGAGGGGTGTGTGACTACTATAGGCTGTATGAGCCTAGCTGGGA 10944  
|||||  
QY 1244 GGTGGATGGCC 1254  
|||||  
Db 10945 GGTGGATGGCC 10955  
|||||

RESULT 9  
HS257120/c 114846 bp DNA linear PRI 22-NOV-2001  
LOCUS  
DEFINITION Human DNA sequence from clone RPI-257120 on chromosome  
22q13.1-13.2, complete sequence.  
ACCESSION AL021878  
VERSION AL021878.2 GI:17065905  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

1 (bases 1 to 114846)  
Bridgeman,A.  
Direct Submission  
Submitted (22-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Nov 25, 2001 this sequence version replaced gi:3204432.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C-elegans/wormpep This sequence  
was finished as follows unless otherwise noted: all regions were  
either double-stranded or sequenced with an alternate chemistry or  
covered by high quality data (i.e., phred quality >= 30); an  
attempt was made to resolve all sequencing problems, such as  
compressions and repeats; all regions were covered by at least one  
plasmid subclone or more than one M13 subclone; and the assembly  
was confirmed by restriction digest. This sequence was generated  
from part of bacterial clone contigs of human chromosome 22,  
constructed by the Sanger Centre Chromosome 22 Mapping Group.  
Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr22  
RPI-257120 is from the library RPCI-1 constructed by the group of  
Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pCYPAC2  
This sequence is the entire insert of clone RPI-257120 The true  
right end of clone RPI-18601 is at 20171 in this sequence.

FEATURES

source

Location/Qualifiers

1..114846  
/organism="Homo sapiens"  
/db\_xref="RZPD:RPCIP704I20257"  
/db\_xref="taxon:9606"  
/chromosome="22"  
/map="q13.1-13.2"  
/clone="RPI-257120"  
/clone\_lib="RPCI-1"  
/note="match: GSS: Em:AQ8871153"  
1093..1682  
/note="match: GSS: Em:AQ563517"  
1097..1600  
/note="match: GSS: Em:AQ563517"  
complement(3894..4138)  
/note="match: GSS: Em:AQ140240"  
4148..4647  
/note="match: GSS: Em:AQ684484"  
7162..7366  
/note="match: STS: Em:HS324WC5"  
7357..7396  
/note="20 copies 2 mer ac 82% conserved"  
7369..7551  
/note="match: STS: Em:HS324WC5"  
7372..9283  
/note="CpG island"  
/evidence="not\_experimental"  
9981..10378  
/note="match: GSS: Em:AQ124532"  
13208..14275  
/note="MER11C repeat: matches 1..1057 of consensus"  
16086..16507  
/note="LIMD repeat: matches 1552..1964 of consensus"  
17055..17127  
/note="LIMD repeat: matches 1482..1552 of consensus"  
17434..17688  
/note="LIMD repeat: matches 1222..1482 of consensus"  
19152..19153  
/note="clone RPI-18601"

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

repeat\_region

misc\_feature

misc\_feature

misc\_feature

repeat\_region

repeat\_region

repeat\_region

repeat\_region

variation

```
aa in this entry
deletion"
/replace="aaataa"
19896..19839
/notes="L1P repeat: matches 2416..2459 of consensus"
repeat_region
/notes="L1P repeat: matches 2378..2416 of consensus"
repeat_region
20246..20284
/notes="L1P repeat: matches 2311..2378 of consensus"
repeat_region
20594..20660
/notes="L1P repeat: matches 2294..2311 of consensus"
repeat_region
20965..20982
/notes="L1P repeat: matches 2294..2311 of consensus"
repeat_region
21270..22130
/notes="L1P repeat: matches 1445..2294 of consensus"
repeat_region
22425..22499
/notes="L1P repeat: matches 1370..1445 of consensus"
repeat_region
22502..22994
/notes="L1R10B repeat: matches 1..510 of consensus"
repeat_region
23020..23379
/notes="Bav-int repeat: matches 4243..6625 of consensus"
misc_feature
25398..25662
/notes="match: STS: Em:G07321"
misc_feature
25597..26083
/notes="match: GSS: Em:B13982"
misc_feature
25611..25853
/notes="match: GSS: Em:B14069"
misc_feature
complement(26269..26446)
/notes="match: GSS: Em:AF046780"
misc_feature
26890..26995
/notes="match: GSS: Em:AZ083430"
repeat_region
27147..28076
/notes="HERVI repeat: matches 4455..5408 of consensus"
repeat_region
28386..30853
/notes="HERVI repeat: matches 2098..4455 of consensus"
repeat_region
30843..30912
/notes="HERVI repeat: matches 2098..2163 of consensus"
repeat_region
30902..30971
/notes="HERVI repeat: matches 2098..2163 of consensus"
repeat_region
30961..31556
/notes="HERVI repeat: matches 1561..2163 of consensus"
repeat_region
31873..33456
/notes="HERVI repeat: matches 9..1561 of consensus"
repeat_region
33464..33959
/notes="L1R10B repeat: matches 1..510 of consensus"
repeat_region
33960..34110
/notes="L1PbA repeat: matches 1226..1378 of consensus"
repeat_region
34415..35367
/notes="L1PbA repeat: matches -656..1226 of consensus"
repeat_region
37163..37323
/notes="Harlequin repeat: matches 912..1086 of consensus"
repeat_region
37324..37437
/notes="57 copies 2 mer ga 79% conserved"
misc_feature
complement(join(42379..42591,42916..43112))
/notes="match: STS: Em:G27508"
misc_feature
complement(join(42491..42603,42916..42945))
/notes="match: STS: Em:G43129"
repeat_region
43992..44915
/notes="ptr5 repeat: matches 40..93 of consensus"
repeat_region
44916..45564
/notes="L1R12 repeat: matches 5..671 of consensus"
misc_feature
48350..49467
/notes="CpG island"
/notes="not_experimental"
repeat_region
54417..54450
/notes="17 copies 2 mer ca 100% conserved"
repeat_region
54578..55083
/notes="MER77 repeat: matches 30..587 of consensus"
misc_feature
58051..59086
/notes="CpG island"
/notes="not_experimental"
misc_feature
63436..65550
/notes="match: GSS: Em:AQ035975"
repeat_region
65916..65961
/notes="23 copies 2 mer tt 82% conserved"
```

```
misc_feature
complement(65929..66250)
/notes="match: STS: Em:G27630"
repeat_region
66514..66549
/notes="18 copies 2 mer aa 80% conserved"
misc_feature
66530..66802
/notes="match: GSS: Em:AQ617313"
misc_feature
66577..66930
/notes="match: GSS: Em:AQ084812"
misc_feature
66603..66893
/notes="match: GSS: Em:AQ044403"
repeat_region
66974..67017
/notes="22 copies 2 mer cc 75% conserved"
misc_feature
complement(69196..69374)
/notes="match: GSS: Em:B14383"
misc_feature
69710..70103
/notes="match: GSS: Em:AQ181535"
misc_feature
complement(78378..78961)
/notes="match: GSS: Em:AQ389013"
misc_feature
complement(78421..78533)
/notes="match: GSS: Em:AQ042556"
misc_feature
complement(78451..78977)
/notes="match: GSS: Em:AQ533248"
misc_feature
complement(81846..82274)
/notes="match: GSS: Em:AQ618257"
repeat_region
82587..83003
/notes="L1MA5A repeat: matches 5857..6292 of consensus"
repeat_region
83609..83702
/notes="MER21B repeat: matches 703..787 of consensus"
repeat_region
83697..83797
/notes="MER21B repeat: matches 686..794 of consensus"
repeat_region
83910..84357
/notes="MER21B repeat: matches 76..540 of consensus"
repeat_region
85375..85430
/notes="28 copies 2 mer aa 76% conserved"
misc_feature
complement(90629..90900)
/notes="match: GSS: Em:B13983"
repeat_region
96343..96364
/notes="11 copies 2 mer ac 100% conserved"
misc_feature
99274..99444
/notes="Other . Weak data"
repeat_region
104028..104091
/notes="32 copies 2 mer aa 67% conserved"
misc_feature
complement(107923..108378)
/notes="match: GSS: Em:AQ572846"
```

Query Match 16.0%; Score 269; DB 9; Length 114846;  
Best Local Similarity 99.5%; Pred. No. 1.4e-128;  
Matches 369; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 884 AGTGTGAGTGTGAGTGTGTATGTGTGAATATGTTGTGTGGTGATTTCTGCT 943
|||||
Db 51089 AGTGTGAGTGTGAGTGTGTATGTGTGAATATGTTGTGTGGTGATTTCTGCT 51030

QY 944 GTGTATCTGTCCCGCAAGTGTGAACAAGTGGACAAGTGTCTGGAGTGGACAAGAGA 1003
|||||
Db 51029 GTGTATCTGTCCCTGCAAGTGTGAACAAGTGGACAAGTGTCTGGAGTGGACAAGAGA 50970

QY 1004 TCTGTSCACCATCAGGTGTGTCATAGCGTCTGCCATGTCGAAGTGTCAAGTGTGAAGTG 1063
|||||
Db 50969 TCTGTSCACCATCAGGTGTGTCATAGCGTCTGCCATGTCGAAGTGTCAAGTGTGAAGTG 50910

QY 1064 AAGGACCAAGGCCCATGATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGTAAGTGCC 1123
|||||
Db 50909 AAGGACCAAGGCCCATGATGCCACTCATCATCAGGAGCTCTAAGGCCCCAGGTAAGTGCC 50850

QY 1124 AGTGACAGATAGGGTGTCTGAAGTCACTCTCGAGTGGCAGTGGGGTGAAGGAAAGG 1183
|||||
Db 50849 AGTGACAGATAGGGTGTCTGAAGTCACTCTCGAGTGGCAGTGGGGTGAAGGAAAGG 50790

QY 1184 CAAGCCCATGTTCTGGAGGAGGGTTGTGACTACATTAGGGTGTATGACCTAGCTGGGA 1243
|||||
Db 50789 CAAGTCAATGTTCTGGAGGAGGGTTGTGACTACATTAGGGTGTATGACCTAGCTGGGA 50730
```



```

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1620 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTCTGCTCCTGGTG 1679
|||||
Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTCTGCTCCTGGTG 60

QY 1680 G 1680
Db 61 G 61

RESULT 13
E10867
LOCUS E10867 1494 bp RNA linear PAT 29-SEP-1997
DEFINITION cdna encoding human cytochrome P450.
ACCESSION E10867
VERSION E10867.1 GI:22027961
KEYWORDS JP 1996056695-A/16.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AYashi.K., Sakaki.T., Yabusaki.Y., Komai.K., Kaneko.H. and
Nakatsuka.I.
METHOD FOR EVALUATING SAFETY
Patent: JP 1996056695-A 16 05-MAR-1996;
SUMITOMO CHEM CO LTD
OS Homo sapiens (human)
PN JP 1996056695-A/16
PD 05-MAR-1996
PF 15-JUL-1994 JP 1994164184
PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR
17-JUN-1994 JP 94P 136053
PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI
KOICHIRO.
PI KANEKO HIDEO, NAKATSUKA IWA0
PC C1201/02,C12M1/34,C12Q1/26;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1494
FT CD5 /organism='Homo sapiens'
FT CDS 1..1494
FT /product='human cytochrome P450 2D6'.
FEATURES
source
1..1494
Location/Qualifiers
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 248 a 508 c 446 g 292 t
ORIGIN
Query Match 3.6%; Score 61; DB 6; Length 1494;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1620 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTCTGCTCCTGGTG 1679
|||||
Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTCTGCTCCTGGTG 60

QY 1680 G 1680
Db 61 G 61

RESULT 14
E10868
LOCUS E10868 1494 bp RNA linear PAT 29-SEP-1997
DEFINITION cdna encoding human cytochrome P450.
ACCESSION E10868

```

```

E10868.1 GI:22027962
JP 1996056695-A/17.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AYashi.K., Sakaki.T., Yabusaki.Y., Komai.K., Kaneko.H. and
Nakatsuka.I.
METHOD FOR EVALUATING SAFETY
Patent: JP 1996056695-A 17 05-MAR-1996;
SUMITOMO CHEM CO LTD
OS Homo sapiens (human)
PN JP 1996056695-A/17
PD 05-MAR-1996
PF 15-JUL-1994 JP 1994164184
PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR
17-JUN-1994 JP 94P 136053
PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI
KOICHIRO.
PI KANEKO HIDEO, NAKATSUKA IWA0
PC C1201/02,C12M1/34,C12Q1/26;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..1494
FT CD5 /organism='Homo sapiens'
FT CDS 1..1494
FT /product='human cytochrome P450 2D6'.
FEATURES
source
1..1494
Location/Qualifiers
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 248 a 509 c 446 g 291 t
ORIGIN
Query Match 3.6%; Score 61; DB 6; Length 1494;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1620 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTCTGCTCCTGGTG 1679
|||||
Db 1 ATGGGGCTAGAGCACTGGTGGCCCTGGCCGTGATAGTGGCCATCTCTGCTCCTGGTG 60

QY 1680 G 1680
Db 61 G 61

RESULT 15
E10869
LOCUS E10869 1494 bp RNA linear PAT 29-SEP-1997
DEFINITION cdna encoding human cytochrome P450.
ACCESSION E10869
VERSION E10869.1 GI:22027963
KEYWORDS JP 1996056695-A/18.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AYashi.K., Sakaki.T., Yabusaki.Y., Komai.K., Kaneko.H. and
Nakatsuka.I.
METHOD FOR EVALUATING SAFETY
Patent: JP 1996056695-A 18 05-MAR-1996;
SUMITOMO CHEM CO LTD
OS Homo sapiens (human)
PN JP 1996056695-A/18
PD 05-MAR-1996
PF 15-JUL-1994 JP 1994164184

```

PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR  
17-JUN-1994 JP 94P 136053  
PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI  
KOICHIRO,  
PI KANEKO HIDEO, NAKATSUKA IWAO  
PC C12Q1/02.C12M1/34.C12Q1/26;  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No; Location/Qualifiers  
FH Key  
FT source 1. .1494  
FT CDS /organism="Homo sapiens"  
FT /db\_xref="taxon:9606"  
FT /product="human cytochrome P450 2D6".  
FEATURES  
source  
Location/Qualifiers  
1. .1494  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 248 a 508 c 447 g 291 t  
ORIGIN

Query Match 3.6%; Score 61; DB 6; Length 1494;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 1679  
|||||  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680  
|  
Db 61 G 61

RESULT 16  
LOCUS AR084365 1494 bp RNA linear PAT 29-SEP-1997  
DEFINITION cDNA encoding human cytochrome P450.  
ACCESSION E10870  
VERSION E10870.1 GI:22027964  
KEYWORDS JP 1996056695-A/19.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Hayashi,K., Sakaki,T., Yabusaki,Y., Komai,K., Kaneko,H. and  
Nakatsuka,I.  
TITLE METHOD FOR EVALUATING SAFETY  
JOURNAL Patent: JP 1996056695-A 19 05-MAR-1996;  
SUMITOMO CHEM CO LTD  
COMMENT OS Homo sapiens (human)  
PN JP 1996056695-A/19  
PD 05-MAR-1996  
PF 15-JUL-1994 JP 1994164184  
PR 20-JUL-1993 JP 93P 201120, 30-JUL-1993 JP 93P 208279, PR  
17-JUN-1994 JP 94P 136053  
PI HAYASHI KOJI, SAKAKI TOSHIYUKI, YABUSAKI YOSHIYASU, PI KOMAI  
KOICHIRO,  
PI KANEKO HIDEO, NAKATSUKA IWAO  
PC C12Q1/02.C12M1/34.C12Q1/26;  
CC strandedness: Double;  
CC topology: Linear;  
CC hypothetical: No;  
CC anti-sense: No; Location/Qualifiers  
FH Key  
FT source 1. .1494  
FT CDS /organism="Homo sapiens"  
FT /product="human cytochrome P450 2D6".

FEATURES  
source  
Location/Qualifiers  
1. .1494  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 248 a 507 c 447 g 292 t  
ORIGIN

Query Match 3.6%; Score 61; DB 6; Length 1494;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 1679  
|||||  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680  
|  
Db 61 G 61

RESULT 17  
LOCUS AR084365 1566 bp DNA linear PAT 01-SEP-2000  
DEFINITION Sequence 13 from patent US 5981174.  
ACCESSION AR084365  
VERSION AR084365.1 GI:10011136  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1566)  
AUTHORS Wolf,C.Roland., Miles,J.Stephen., Spurr,N.Kay. and Gough,A.Charles.  
TITLE Genetic assay  
JOURNAL Patent: US 5981174-A 13 09-NOV-1999;  
FEATURES  
source  
Location/Qualifiers  
1. .1566  
/organism="unknown"  
BASE COUNT 268 a 526 c 464 g 308 t  
ORIGIN

Query Match 3.6%; Score 61; DB 6; Length 1566;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 1679  
|||||  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680  
|  
Db 61 G 61

RESULT 18  
LOCUS AR084374 1566 bp DNA linear PAT 01-SEP-2000  
DEFINITION Sequence 22 from patent US 5981174.  
ACCESSION AR084374  
VERSION AR084374.1 GI:10011145  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1566)  
AUTHORS Wolf,C.Roland., Miles,J.Stephen., Spurr,N.Kay. and Gough,A.Charles.  
TITLE Genetic assay  
JOURNAL Patent: US 5981174-A 22 09-NOV-1999;  
FEATURES  
source  
Location/Qualifiers  
1. .1566  
/organism="unknown"  
BASE COUNT 269 a 531 c 461 g 305 t  
ORIGIN

```
Query Match          3.6%; Score 61; DB 6; Length 1566;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTCCCTGCGCGGTAGTGGCCATCTTCCTGCTCCTGGTG 1679
      |||||||
Db 1 ATGGGGCTAGAACACTGGTCCCTGCGCGGTAGTGGCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680
      |
Db 61 G 61

RESULT 19
HSDB1
LOCUS      Homo sapiens mRNA for cytochrome P450 db1.          PRI 15-JUN-2000
DEFINITION
ACCESSION  X08006 Y00300
VERSION    X08006.1 GI:30450
KEYWORDS   cytochrome P450; cytochrome P450 db1.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 1567)
REFERENCE  Gonzalez,F.J., Skoda,R.C., Kimura,S., Umeno,M., Zanger,U.M.,
AUTHORS    Nebert,D.W., Gelboin,H.V., Hardwick,J.P. and Meyer,U.A.
TITLE      Characterization of the common genetic defect in humans deficient
            in debrisoquine metabolism
JOURNAL    Nature 331 (6155), 442-446 (1988)
MEDLINE    88122614
PUBMED     3123997
REFERENCE  2 (bases 1 to 1567)
AUTHORS    Nebert,D.W.
TITLE      Direct Submission
JOURNAL    Submitted (30-JUN-1988) Nebert D.W., NIH, Bethesda, Maryland 20892
COMMENT    see also X07618 (variant a), X07619 (variant b) and X07620 (variant
            b').

FEATURES             Location/Qualifiers
     source           1..1567
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="22"
                     /map="22q11.2-qter"
                     /cell_type="hepatocyte"
                     /tissue_type="liver"
                     /clone_lib="lambda gt11"
     CDS               1..1494
                     /codon_start=1
                     /product="cytochrome P450 db1"
                     /protein_id="CAA30807.1"
                     /db_xref="GI:30451"
                     /db_xref="SWISS-PROT:P10635"
                     /translation="MGLEALVPLAVIAIFLLVLDLMHRRQWAAARYPPGLPLPLGLG
                     NLLHVDQNTFYCFDQLRFRGDFVSLQAWTPVVVNLGLAAVREALVTHGEDTADRP
                     PVPITQILGFGPRSGVFLARYGPWQRORFVSVTLRNLGLGKKSLSQWVEEAACL
                     CAAPAHNSGRPRPRLGLLDKAVSNVIASTCTGRFEYDDPRFLRLDLAQGLKEESG
                     FLREVLNAVPLLHLPALAGKVLRFQKAFLTQLDELLTEHRMTWDPAPPRDLTEAFL
                     AEMEKAKNPRESSNDENLRIVVADLFAGVMVTTSTLLAWGLLLMLHPDVORRQVEE
                     IDDVIGQVRPEMGDQAHMPYTTAVIHEVQREGDIVPLGTHMTSRDIEVQGFRIKPG
                     TLLTNLSVLKDEAVWEKPPFHPHEFLDAQGHVKEPEAFLPFSAGRRACLGEPFLAR
                     MELELFFSLLOHFSFSVPTGQPRFSHHGVFAFLVSPSPYELCAVPR"
     polyA_signal      1546..1550
                     /note="putative"
     polyA_site        1567
                     /note="polyA site"
BASE COUNT          267 a 527 c 465 g 308 t
ORIGIN
Query Match          3.6%; Score 61; DB 9; Length 1567;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTCCCTGCGCGGTAGTGGCCATCTTCCTGCTCCTGGTG 1679
      |||||||
Db 1 ATGGGGCTAGAACACTGGTCCCTGCGCGGTAGTGGCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680
      |
Db 61 G 61

RESULT 21
HUMCYPDB1
LOCUS      Human cytochrome P450 db1 mRNA, complete cds.
DEFINITION
ACCESSION  M20403 M19697
VERSION    M20403.1 GI:181349
KEYWORDS   cytochrome P450; debrisoquine 4-hydroxylase.
SOURCE     Human hepatocyte, cDNA to mRNA.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 1567)
REFERENCE  Gonzalez,F.J., Vilbois,F., Hardwick,J.P., McBride,O.W.,
AUTHORS    Nebert,D.W., Gelboin,H.V. and Meyer,U.A.
TITLE      Human debrisoquine 4-hydroxylase (P450IID1): cDNA and deduced amino
            acid sequence and assignment of the CYP2D locus to chromosome 22
JOURNAL    Genomics 2 (2), 174-179 (1988)
MEDLINE    88334109
PUBMED     3410476
COMMENT    Draft entry and printed copy of sequence for [1] kindly provided by
            D.W.Nebert, 15-JUL-1988.

FEATURES             Location/Qualifiers
     source           1..1567
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="22q11.2-qter"
     gene             1..1567
                     /gene="CYP2D"
     mRNA             <1..1567
                     /gene="CYP2D"
                     /product="CY P450 db1 mRNA"
                     /db_xref="GI:181350"
                     /protein_id="AAA52153.1"
                     /note="debrisoquine 4-hydroxylase"
     CDS               1..1494
                     /codon_start=1
                     /db_xref="GI:181350"
                     /db_xref="GDB:G00-119-832"
                     /translation="MGLEALVPLAVIAIFLLVLDLMHRRQWAAARYPPGLPLPLGLG
                     NLLHVDQNTFYCFDQLRFRGDFVSLQAWTPVVVNLGLAAVREALVTHGEDTADRP
                     PVPITQILGFGPRSGVFLARYGPWQRORFVSVTLRNLGLGKKSLSQWVEEAACL
                     CAAPAHNSGRPRPRLGLLDKAVSNVIASTCTGRFEYDDPRFLRLDLAQGLKEESG
                     FLREVLNAVPLLHLPALAGKVLRFQKAFLTQLDELLTEHRMTWDPAPPRDLTEAFL
                     AEMEKAKNPRESSNDENLRIVVADLFAGVMVTTSTLLAWGLLLMLHPDVORRQVEE
                     IDDVIGQVRPEMGDQAHMPYTTAVIHEVQREGDIVPLGTHMTSRDIEVQGFRIKPG
                     TLLTNLSVLKDEAVWEKPPFHPHEFLDAQGHVKEPEAFLPFSAGRRACLGEPFLAR
                     MELELFFSLLOHFSFSVPTGQPRFSHHGVFAFLVSPSPYELCAVPR"
BASE COUNT          267 a 527 c 465 g 308 t
ORIGIN
Query Match          3.6%; Score 61; DB 9; Length 1567;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTCCCTGCGCGGTAGTGGCCATCTTCCTGCTCCTGGTG 1679
      |||||||
Db 1 ATGGGGCTAGAACACTGGTCCCTGCGCGGTAGTGGCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680
      |
Db 61 G 61

RESULT 21
```

A20907 LOCUS A20907 1568 bp mRNA linear PAT 08-JUN-1994  
DEFINITION debrisouquine hydroxylase.  
ACCESSION A20907  
VERSION A20907.1 GI:583441  
KEYWORDS synthetic construct.  
SOURCE synthetic construct.  
ORGANISM artificial sequences.  
FEATURES  
    source  
        Location/Qualifiers  
            1..1568  
            /organism="synthetic construct"  
            /db\_xref="taxon:32630"  
            1..1494  
            /codon\_start=1  
            /transl\_table=11  
            /product="debrisouquine hydroxylase"  
            /protein\_id="CAA01522.1"  
            /db\_xref="GI:583442"  
            /translation="MGLEALVPLAVIVAIIFLLLVLMHRRQRWAARYPPGPIPLPLGLG  
            NLLHVDQNTPYCFQDLRRRFGDVSFSLQAWTFVVLNGLAAVREALVTHGEDTADRP  
            PVPITQILGFGSPQSVFLARYGPWQRREFSVTLRNGLGKKSLEQWTEEAACL  
            CAAFANHSGRPRPGLLDKAVSNVIASLTGRRREYDDPRFLRLDLAQEGLKESG  
            FLREVLNAPVLLHIPALAGKVLRFQKAFLTQLDELLEHRTWDPQAPPRDLTEAFL  
            AMEKAKNPESFNDENKRIYVADLFSAGWITSTLAWGLLMLLHPDVPQVSNRR  
            STTWIGQVRPMPGQAHMPYITAVIHEVQREGDIVPLGTMHTSDIEVGFRIPKG  
            TTLITNLSSVLKDEAVWEKPFREPHPEHFLDAQGHFVKPEAFLPFSAGRACILGELPAR  
            MELFLFFTSLLQHSFVSPTGQPRPSHHGVFAFLVSPSPYELCAVPR"  
BASE COUNT 268 a 527 c 465 g 308 t  
ORIGIN  
  
Query Match 3.6%; Score 61; DB 6; Length 1568;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTCCGCTCCTGGTG 1679  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTCCGCTCCTGGTG 60  
  
QY 1680 G 1680  
Db 61 G 61  
  
RESULT 22  
LOCUS AR084372 1568 bp DNA linear PAT 01-SEP-2000  
DEFINITION Sequence 20 from patent US 5981174.  
ACCESSION AR084372  
VERSION AR084372.1 GI:10011143  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Wolf,C.Roland., Miles,J.Stephen., Spurr,N.Kay. and Gough,A.Charles.  
TITLE Genetic assay  
JOURNAL Patent: US 5981174-A 20 09-NOV-1999;  
FEATURES  
    source  
        Location/Qualifiers  
            1..1568  
            /organism="unknown"  
BASE COUNT 267 a 527 c 464 g 310 t  
ORIGIN  
  
Query Match 3.6%; Score 61; DB 6; Length 1568;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTCCGCTCCTGGTG 1679  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTCCGCTCCTGGTG 60  
  
QY 1680 G 1680  
Db 61 G 1680

Db 61 G 61  
  
RESULT 23  
LOCUS AR084373 1571 bp DNA linear PAT 01-SEP-2000  
DEFINITION Sequence 21 from patent US 5981174.  
ACCESSION AR084373  
VERSION AR084373.1 GI:10011144  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1571)  
AUTHORS Wolf,C.Roland., Miles,J.Stephen., Spurr,N.Kay. and Gough,A.Charles.  
TITLE Genetic assay  
JOURNAL Patent: US 5981174-A 21 09-NOV-1999;  
FEATURES  
    source  
        Location/Qualifiers  
            1..1571  
            /organism="unknown"  
BASE COUNT 270 a 529 c 468 g 304 t  
ORIGIN  
  
Query Match 3.6%; Score 61; DB 6; Length 1571;  
Best Local Similarity 100.0%; Pred. No. 1.9e-20;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTCCGCTCCTGGTG 1679  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGTGATAGTGCCCATCTCCGCTCCTGGTG 60  
  
QY 1680 G 1680  
Db 61 G 61  
  
RESULT 24  
LOCUS AC019129 172611 bp DNA linear PRI 09-JAN-2002  
DEFINITION Homo sapiens BAC clone RP11-559M23 from 2, complete sequence.  
ACCESSION AC019129  
VERSION AC019129.8 GI:16077059  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 172611)  
AUTHORS Sulston,J.E. and Waterston,R.  
TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
MEDLINE 99063792  
PUBMED 9847074  
REFERENCE 2 (bases 1 to 172611)  
AUTHORS Trani,L., Abbott,A. and Creason,K.  
TITLE The sequence of Homo sapiens BAC clone RP11-559M23  
JOURNAL Unpublished (2001)  
REFERENCE 3 (bases 1 to 172611)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (30-DEC-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
REFERENCE 4 (bases 1 to 172611)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (12-OCT-2001) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
REFERENCE 5 (bases 1 to 172611)  
AUTHORS Waterston,R.  
TITLE Direct Submission

## JOURNAL

Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Oct 12, 2001 this sequence version replaced gi:156224993.

## COMMENT

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)  
----- Summary Statistics  
Center project name: H\_NH0559M23  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

## SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoeawa, K., Woon, P.Y., Zhao, B., Frenken, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)  
VECTOR: pBACE3.6

## NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is AC087053; the clone sequenced to the right is RP11-2415, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-559M23; actual end is at base position 17064 of RP11-2415.

Single plasmid coverage exists between bases 30202 to 30267.

## FEATURES

source

Location/Qualifiers

1. .172611  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2"  
/clone="RP11-559M23"  
/clone\_lib="RPCI-11"

repeat\_region

1. .1439

/rpt\_family="ERV1"

repeat\_region

1440. .1738

/rpt\_family="Alu"

repeat\_region

1739. .1964

/rpt\_family="ERV1"

repeat\_region

1977. .2082

/rpt\_family="ERV1"

repeat\_region

2087. .2218

/rpt\_family="Alu"

repeat\_region

3209. .3259

/rpt\_family="(TTTGG)n"

repeat\_region

3230. .3551

/rpt\_family="Alu"

repeat\_region

3406. .3426

/rpt\_family="AT-rich"

repeat\_region

3554. .3580

/rpt\_family="AT-rich"  
4188. .4453  
/rpt\_family="Alu"  
4453. .4634  
/rpt\_family="GA-rich"  
4672. .4976  
/rpt\_family="Alu"  
4953. .4976  
/rpt\_family="(A)n"  
5162. .5399  
/rpt\_family="L2"  
5430. .5588  
/rpt\_family="MER1\_type?"  
5559. .5594  
/rpt\_family="AT-rich"  
5608. .5628  
/rpt\_family="(TTTA)n"  
5611. .5744  
/rpt\_family="Alu"  
6207. .6265  
/rpt\_family="MER1\_type?"  
6338. .6625  
/rpt\_family="MER1\_type?"  
6543. .6579  
/rpt\_family="AT-rich"  
6671. .6732  
/rpt\_family="Alu"  
6755. .7031  
/rpt\_family="ERV1"  
7059. .7098  
/rpt\_family="ERV1"  
7096. .7166  
/rpt\_family="MER1\_type?"  
7428. .7712  
/rpt\_family="Alu"  
7720. .8056  
/rpt\_family="L2"  
8429. .8537  
/rpt\_family="Alu"  
8708. .8756  
/rpt\_family="L2"  
9394. .9432  
/rpt\_family="(TTTA)n"  
9417. .9708  
/rpt\_family="Alu"  
9710. .9730  
/rpt\_family="(TTTA)n"  
9712. .10012  
/rpt\_family="Alu"  
10226. .10618  
/rpt\_family="ERV1"  
11969. .12201  
/rpt\_family="MIR"  
12637. .12795  
/rpt\_family="MIR"  
12930. .13232  
/rpt\_family="Alu"  
13351. .13421  
/rpt\_family="MER103"  
13422. .13717  
/rpt\_family="Alu"  
13718. .13756  
/rpt\_family="MER103"  
13831. .14034  
/rpt\_family="MIR"  
14518. .14566  
/rpt\_family="T-rich"  
15603. .15624  
/rpt\_family="AT-rich"  
16436. .16680  
/rpt\_family="L2"  
16672. .16710  
/rpt\_family="(T)n"



```
repeat_region 16681..16978 /rpt_family="Alu"
repeat_region 16979..17051 /rpt_family="L2"
repeat_region 17123..17255 /rpt_family="Alu"
repeat_region 17238..17264 /rpt_family="Alu"
repeat_region 17347..17732 /rpt_family="AT-rich"
repeat_region 17667..17864 /rpt_family="L2"
repeat_region 17767..17864 /rpt_family="L2"
repeat_region 18006..18289 /rpt_family="MER53"
repeat_region 18361..18553 /rpt_family="MALR"
repeat_region 18716..18742 /rpt_family="L2"
repeat_region 19412..19455 /rpt_family="(TTTG)n"
repeat_region 19412..19455 /rpt_family="(TG)n"

Query Match          3.6%; Score 61; DB 9; Length 172611;
Best Local Similarity 100.0%; Pred. No. 8.4e-21;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 AGTGGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAAC 175
|||||
Db 137700 AGTGGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAAC 137641

QY 176 C 176
Db 137640 C 137640

RESULT 25
AC007551/C
LOCUS AC007551 155401 bp DNA linear PRI 13-MAY-1999
DEFINITION Homo sapiens clone RP11-239C9 from 7p14-15, complete sequence.
ACCESSION AC007551
VERSION AC007551.1 GI:4809347
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 155401)
AUTHORS Iadonato,S.P., Yu,J., Wong,G.K.-S., Magness,C.L., Green,E.D.,
Green,P. and Olson,M.V.
Large-scale MCD Mapping and Sequencing of Human Chromosome 7
Unpublished
2 (bases 1 to 155401)
Bubb,K.L. and Desmarais,C.L.
Direct Submission
Submitted (13-MAY-1999) Human Genome Center, University of
Washington, Box 352145, Seattle, WA 98195, USA
Overlapping Sequences:
5': UWGC: djs74
3': mapping in progress
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.

Double stranded (DS) coverage: 100.0%
DS or two chemistry coverage: 100.0%
Single stranded regions: 0
-----
```

Sequence Validation:  
This sequence has been validated by Multiple Complete Digest  
fingerprinting. Comparison of the experimentally derived digest  
fragments with sequence-predicted fragments is given below.  
Small fragments below a variable cutoff (approximately 400-600)  
are not resolved in the fingerprint and hence do not appear  
in the table. There are no significant remaining discrepancies  
between the experimental and predicted values. Uniquely ordered  
fragment groups are separated by dashed lines.

FP	Seq	FP	Seq	FP	Seq
BgIII	ECORI	BgIII	ECORI	BgIII	Seq
-----	-----	-----	-----	-----	-----
2650.00	2667.00	1122.00	1150.00	669.00	652.00
-----	-----	-----	-----	-----	-----
1258.00	1279.00	7714.00	7740.00	2239.00	2271.00
-----	-----	-----	-----	-----	-----
4794.00	4702.00	10502.00	10436.00	7685.00	7510.00
-----	-----	-----	-----	-----	-----
6880.00	6992.00	2254.00	2303.00	6318.00	6339.00
-----	-----	-----	-----	-----	-----
3779.00	3753.00	1494.00	1471.00	5903.00	5951.00
-----	-----	-----	-----	-----	-----
5734.00	5823.00	4522.00	4501.00	8292.00	8250.00
-----	-----	-----	-----	-----	-----
4575.00	4581.00	1714.00	1721.00	1967.00	1930.00
-----	-----	-----	-----	-----	-----
1060.00	1091.00	8383.00	8252.00	18377.00	18231.00
-----	-----	-----	-----	-----	-----
5734.00	5623.00	2668.00	2682.00	7685.00	7642.00
-----	-----	-----	-----	-----	-----
3779.00	3800.00	1038.00	1024.00	16295.00	16345.00
-----	-----	-----	-----	-----	-----
10546.00	10571.00	9711.00	9647.00	5590.00	5573.00
-----	-----	-----	-----	-----	-----
5305.00	5285.00	532.00	531.00	2922.00	2920.00
-----	-----	-----	-----	-----	-----
1632.00	1618.00	558.00	585.00	7685.00	7724.00
-----	-----	-----	-----	-----	-----
943.00	929.00	17406.00	17191.00	669.00	638.00
-----	-----	-----	-----	-----	-----
1169.00	1162.00	769.00	750.00	3630.00	3665.00
-----	-----	-----	-----	-----	-----
6880.00	6999.00	5002.00	5016.00	810.00	786.00
-----	-----	-----	-----	-----	-----
3411.00	3350.00	1554.00	1544.00	4577.00	4548.00
-----	-----	-----	-----	-----	-----
6114.00	6187.00	1494.00	1518.00	716.00	714.00
-----	-----	-----	-----	-----	-----
1357.00	1353.00	1432.00	1414.00	3173.00	3143.00
-----	-----	-----	-----	-----	-----
4273.00	4302.00	614.00	615.00	969.00	957.00
-----	-----	-----	-----	-----	-----
7295.00	7354.00	3074.00	3068.00	810.00	813.00
-----	-----	-----	-----	-----	-----
6446.00	6484.00	1122.00	1111.00	982.00	1013.00
-----	-----	-----	-----	-----	-----
6679.00	6657.00	2419.00	2361.00	9745.00	9674.00
-----	-----	-----	-----	-----	-----
1060.00	1012.00	2783.00	2814.00	2128.00	2109.00
-----	-----	-----	-----	-----	-----
1169.00	1146.00	7714.00	7592.00	2128.00	2151.00
-----	-----	-----	-----	-----	-----
3411.00	3380.00	9024.00	9070.00	1342.00	1330.00
-----	-----	-----	-----	-----	-----
3687.00	3703.00	6602.00	6603.00	3037.00	3017.00
-----	-----	-----	-----	-----	-----
4273.00	4288.00	5181.00	5192.00	982.00	958.00
-----	-----	-----	-----	-----	-----
5136.00	5089.00	1987.00	1992.00	3425.00	3366.00
-----	-----	-----	-----	-----	-----
1460.00	1457.00	10775.00	10621.00	1500.00	1519.00
-----	-----	-----	-----	-----	-----
2060.00	2040.00	6378.00	6313.00	9348.00	9211.00
-----	-----	-----	-----	-----	-----

```

1838.00 1819.00 1392.00 1369.00 3281.00 3340.00
4794.00 4784.00 2419.00 2403.00 2454.00 2447.00
1060.00 1048.00 2419.00 2423.00 524.00 516.00
12034.00 11986.00 1122.00 1111.00 2454.00 2421.00
4794.00 4752.00 5002.00 4973.00 13475.00
4440.00 4347.00 4522.00 4451.00
605.00 583.00
5136.00
2339.00
2060.00

```

Unmatched fragments are due to vector-containing sequences. We expected 3, 1 and 1-2 such fragments in the BglII, EcoRI, and NsiI enzyme domains, respectively. The sum of these fragments can be calculated from the sequence of the BAC clone and, in this clone, should be 9458, 8782, and 13404 for the three enzyme domains.

#### FEATURES

```

source
1. 155401
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7p14-15"
/clone="RP11-239C9"
/cell_line="Wale Blood"
/clone_lib="RPC-11 Human Male BAC library"
complement(28..79)
/rpt_family="MIR"
complement(1007..1197)
/rpt_family="MLT1"
complement(1361..1452)
/rpt_family="MLT1"
complement(2641..2920)
/rpt_family="Alu"
complement(4878..5175)
/rpt_family="Alu"
complement(6199..6487)
/rpt_family="Alu"
complement(6805..7073)
/rpt_family="Alu"
complement(7895..8179)
/rpt_family="Alu"
8298..8602
/rpt_family="Alu"
9907..10185
/rpt_family="Alu"
11405..11901
/rpt_family="L1"
12136..12399
/rpt_family="Alu"
complement(15140..15429)
/rpt_family="Alu"
17437..17727
/rpt_family="Alu"
complement(20444..20581)
/rpt_family="Alu"
complement(22334..22414)
/rpt_family="MLT1"
complement(22728..22868)
/rpt_family="MIR"
23222..23481
/rpt_family="L1"
23871..23931
/rpt_family="MIR"
24318..24857
/rpt_family="LTR1"
complement(25363..25680)

```

```

/rpt_family="Alu"
30447..30750
/rpt_family="Alu"
30767..30870
/rpt_family="MIR"
30875..31174
/rpt_family="Alu"
36417..36660
/rpt_family="L1"
36878..37160
/rpt_family="Alu"
37585..37863
/rpt_family="Alu"
38028..38321
/rpt_family="Alu"
complement(38886..44355)
/rpt_family="L1"

```

Query Match 3.6%; Score 60; DB 9; Length 155401;  
Best Local Similarity 100.0%; Pred. No. 2.8e-20;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 TAATCCAGCAGCTTTGGAGCCTGAGGTGGTGGATCACCTGAAGTCAGGAGTTCAAGAC 151  
|||||  
Db 69498 TAATCCAGCAGCTTTGGAGCCTGAGGTGGTGGATCACCTGAAGTCAGGAGTTCAAGAC 69439

RESULT 26  
AC073278

LOCUS AC073278 200807 bp DNA linear PRI 09-JAN-2002  
DEFINITION Homo sapiens BAC clone RP11-661L8 from 7, complete sequence.  
ACCESSION AC073278  
VERSION AC073278.9 GI:15145619

KEYWORDS  
SOURCE HTG.

ORGANISM  
Homo sapiens.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 200807)  
Sulston,J.E. and Waterston,R.

Toward a complete human genome sequence  
Genome Res. 8 (11), 1097-1108 (1998)

99063792  
PUBMED 9847074

REFERENCE 2 (bases 1 to 200807)

Harris,A., Abbott,A., Boyer,E. and Elliott,G.

The sequence of Homo sapiens BAC clone RP11-661L8

Unpublished (2001)

REFERENCE 3 (bases 1 to 200807)

Waterston,R.H.

Direct Submission

TITLE

JOURNAL

Submitted (12-JUN-2000) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

REFERENCE 4 (bases 1 to 200807)

Waterston,R.H.

Direct Submission

TITLE

JOURNAL

Submitted (09-AUG-2001) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

REFERENCE 5 (bases 1 to 200807)

Waterston,R.

Direct Submission

TITLE

JOURNAL

Submitted (09-JAN-2002) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

On Aug 9, 2001 this sequence version replaced gi:14476026.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)

----- Summary Statistics

Center project name: H\_NH0661L08

-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tanen, M., Catanese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-627E12; the clone sequenced to the right is RP5-1102A12, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-661L8.

Polymorphisms exist between AC0732278 and AC004963.

#### FEATURES

##### Source

Location/Qualifiers  
1..200807  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosomes="7"  
/map="7"  
/clone\_lib="RPCI-11"

38..98  
/rpt\_family="CR1"  
645..935  
/rpt\_family="Alu"  
1113..1410  
/rpt\_family="Alu"  
1858..2033  
/rpt\_family="Alu"  
2202..2308  
/rpt\_family="Alu"  
2392..2428  
/rpt\_family="L2"  
2615..2768  
/rpt\_family="CT-rich"  
2752..3035  
/rpt\_family="Alu"  
3036..3197  
/rpt\_family="MIR"  
3357..3458  
/rpt\_family="C-rich"  
3456..3503  
/rpt\_family="(TTTTC)n"  
3480..3784  
/rpt\_family="Alu"

repeat\_region 3982..4295  
/rpt\_family="Alu"  
repeat\_region 4302..4377  
/rpt\_family="L1"  
repeat\_region 4518..4792  
/rpt\_family="Alu"  
repeat\_region 4791..4812  
/rpt\_family="(CA)n"  
repeat\_region 4852..4934  
/rpt\_family="L1"  
repeat\_region 4927..5144  
/rpt\_family="L1"  
repeat\_region 5145..5427  
/rpt\_family="Alu"  
repeat\_region 5428..5796  
/rpt\_family="L1"  
repeat\_region 5882..5915  
/rpt\_family="Alu"  
repeat\_region 6061..6394  
/rpt\_family="L1"  
repeat\_region 6384..6430  
/rpt\_family="(TG)n"  
repeat\_region 6430..6684  
/rpt\_family="Alu"  
repeat\_region 6795..6825  
/rpt\_family="(CA)n"  
repeat\_region 6826..6932  
/rpt\_family="(TA)n"  
repeat\_region 6932..7226  
/rpt\_family="Alu"  
repeat\_region 7086..7113  
/rpt\_family="AT-rich"  
repeat\_region 7228..7370  
/rpt\_family="Alu"  
repeat\_region 7231..7254  
/rpt\_family="AT-rich"  
repeat\_region 7574..7599  
/rpt\_family="AT-rich"  
repeat\_region 7868..8008  
/rpt\_family="Alu"  
repeat\_region 7986..8010  
/rpt\_family="AT-rich"  
repeat\_region 8031..8053  
/rpt\_family="BC200"  
repeat\_region 8054..8418  
/rpt\_family="MaLR"  
repeat\_region 8419..8543  
/rpt\_family="BC200"  
repeat\_region 8671..8796  
/rpt\_family="MIR"  
repeat\_region 9175..9452  
/rpt\_family="Alu"  
repeat\_region 9481..9767  
/rpt\_family="Alu"  
repeat\_region 9750..9771  
/rpt\_family="AT-rich"  
repeat\_region 9772..10100  
/rpt\_family="Alu"  
repeat\_region 9900..9936  
/rpt\_family="(CA)n"  
repeat\_region 10075..10113  
/rpt\_family="A-rich"  
repeat\_region 10412..10476  
/rpt\_family="T-rich"  
repeat\_region 10451..10753  
/rpt\_family="Alu"  
repeat\_region 11141..11501  
/rpt\_family="L2"  
repeat\_region 11502..11745  
/rpt\_family="Alu"  
repeat\_region 11577..11600  
/rpt\_family="AT-rich"  
repeat\_region 11746..11768



repeat\_region /rpt\_family="Alu"  
13007.13990  
repeat\_region /rpt\_family="L1"  
14340.14665  
repeat\_region /rpt\_family="ERV1"  
14666.14763  
repeat\_region /rpt\_family="ERV1"  
14821.16006  
repeat\_region /rpt\_family="ERV1"  
16007.16197  
repeat\_region /rpt\_family="MIR"  
16625.16844  
repeat\_region /rpt\_family="MIR"  
17187.17591  
misc\_feature /note="match to EST BG900533 (NID:g14310782)"  
17442.18212  
misc\_feature /note="match to EST BG286769 (NID:g13039938)"  
17715.17816  
repeat\_region /rpt\_family="L2"  
20448.20748  
repeat\_region /rpt\_family="Alu"  
21392.21430  
repeat\_region /rpt\_family="AT-rich"  
21417.21547  
repeat\_region /rpt\_family="Alu"  
21829.21869  
repeat\_region /rpt\_family="(TTTG)n"  
22158.22381  
repeat\_region /rpt\_family="MaLR"  
22382.22933  
repeat\_region /rpt\_family="ERV1"  
22934.23085  
repeat\_region /rpt\_family="MaLR"  
23476.23530  
repeat\_region /rpt\_family="MIR"  
23566.23764  
repeat\_region /rpt\_family="L2"  
23771.23865  
repeat\_region /rpt\_family="MIR"  
23954.23992  
repeat\_region /rpt\_family="(TG)n"  
23991.24674  
repeat\_region /rpt\_family="L1"  
26139.26351  
repeat\_region /rpt\_family="L2"  
26446.26754  
repeat\_region /rpt\_family="Alu"  
27026.27525  
repeat\_region /rpt\_family="MaLR"  
27530.28226  
misc\_feature /note="match to EST BG563093 (NID:g13570745)"  
27550.27936  
misc\_feature /note="match to EST BF817666 (NID:g12154463)"  
28378.28379  
repeat\_region /note="match to EST BG563093 (NID:g13570745)"  
28654.28967  
repeat\_region /rpt\_family="L2"  
28982.29264  
repeat\_region /rpt\_family="Alu"  
31534.31699  
repeat\_region /rpt\_family="MIR"  
32615.32814  
repeat\_region /rpt\_family="MaLR"  
32803.32967  
repeat\_region /rpt\_family="MERL\_type"  
32961.32995  
repeat\_region /rpt\_family="(TTTA)n"  
32968.33270  
repeat\_region /rpt\_family="Alu"  
33271.33303  
repeat\_region /rpt\_family="MERL\_type"  
33304.33606  
repeat\_region /rpt\_family="Alu"

repeat\_region 33586.33605  
repeat\_region /rpt\_family="(CAAAA)n"  
33607.33687  
Query Match 3.5%; Score 58; DB 9; Length 190349;  
Best Local Similarity 100.0%; Pred.No.3e-19;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 77 GGTGGCTCATGCTATATATCCAGCAGCTTTTGGGAGCCTGAGGTGGTGGATCACCTCA 134  
|||||  
Db 126523 GGTGGCTCATGCTATATATCCAGCAGCTTTTGGGAGCCTGAGGTGGTGGATCACCTCA 126466  
RESULT 28  
AL513365/c  
LOCUS AL513365 107415 bp DNA linear PRI 20-JUL-2002  
DEFINITION Human DNA sequence from clone RP3-476K8 on chromosome 1, complete  
sequence.  
ACCESSION AL513365  
VERSION AL513365.27 GI:21953258  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 107415)  
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Lad,H. Catarrhini; Hominidae; Homo.  
TITLE Direct Submission  
JOURNAL Submitted (20-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Jul 24, 2002 this sequence version replaced gi:18564638.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em:, EMBL; Sw:.,  
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP  
database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chrl  
RP3-476K8 is from the library RPCI-3 constructed by the group of  
Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pCYPAC2  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----  
FEATURES  
Source Location/Qualifiers  
1.107415  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RP3-476K8"  
/clone\_lib="RPCI-3"  
BASE COUNT 29399 a 25329 c 25209 g 27478 t  
ORIGIN

```

Query Match          3.3%; Score 56; DB 9; Length 107415;
Best Local Similarity 100.0%; Pred. No. 3.6e-18;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 121 GGTGGATCACCTGAAGTCAGAGTTCAGACTAGGCTGGCCAAACATGTTGTAACCC 176
      |||||||
Db 67252 GGTGGATCACCTGAAGTCAGAGTTCAGACTAGGCTGGCCAAACATGTTGTAACCC 67197

RESULT 29
AL732423/C
LOCUS          AL732423          173519 bp      DNA      linear      HTG 13-AUG-2002
DEFINITION    Homo sapiens chromosome X clone RP11-345L8, *** SEQUENCING IN
PROGRESS ***, in ordered pieces.
ACCESSION     AL732423 AC017058
VERSION       AL732423.11 GI:22265447
KEYWORDS      HTG; HTGS_PHASE3; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 173519)
Direct Submission
Submitted (09-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 15, 2002 this sequence version replaced gi:21738620.
Draft Sequence produced by Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
http://genome.wustl.edu/gsc/index.shtml
----- Genome Center
Center: Wellcome Trust Sanger Institute
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA345L8
----- Summary Statistics
Sequencing program: XGAP4; version 4.5
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-primer-amersham; 24% of reads
Chemistry: Dye-terminator; 6% of reads
Chemistry: Dye-terminator Big Dye; 5% of reads
Consensus quality: 173499 bases at least Q40
Consensus quality: 173519 bases at least Q30
Consensus quality: 173519 bases at least Q20
Insert size: 173519; sum-of-contigs
Insert size: 176928; 4.2% error; agarose-fp
Quality coverage: 11.02x in Q20 bases; sum-of-contigs Quality
coverage: 10.88x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source          Location/Qualifiers
1..173519
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="X"
   /clone="RP11-345L8"
   /clone_lib="RPC1-11.2"
misc_feature    1..173519
   /note="assembly_fragment:03759
clone end:SP6
vector side:right"

BASE COUNT      52534 a 37820 c 37404 g 45761 t
ORIGIN

Query Match          3.3%; Score 56; DB 2; Length 173519;

```

```

Best Local Similarity 100.0%; Pred. No. 3.3e-18;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 GGTGGATCACCTGAAGTCAGAGTTCAGACTAGGCTGGCCAAACATGTTGTAACCC 175
      |||||||
Db 32966 GGTGGATCACCTGAAGTCAGAGTTCAGACTAGGCTGGCCAAACATGTTGTAACCC 32911

RESULT 30
AC080014
LOCUS          AC080014          163111 bp      DNA      linear      PRI 16-MAY-2002
DEFINITION    Homo sapiens 3q BAC RP11-301G23 (Roswell Park Cancer Institute
Human BAC Library) complete sequence.
ACCESSION     AC080014
VERSION       AC080014.13 GI:19033395
KEYWORDS      HTG.
SOURCE        Homo sapiens
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 163111)
Muzny,D.M., Adams,C., Adio-oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,K., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davalila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hoques,M., Holloway,C.,
Hollins,B., Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louiseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mayhew,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Monabbat,K.,
Moore,S., Morgan,M., Morish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savary,G., Scherer,S.,
Scott,G., Shen,H., Shooshitari,N., Sisson,I., Sodergren,E.,
Sonaite,I., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabors,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 163111)
Worley,K.C.
Direct Submission
Submitted (23-SEP-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 163111)
Worley,K.C.
Direct Submission

```

## JOURNAL

Submitted (01-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 163111)  
Worley, K.C.

## REFERENCE

Direct Submission

## TITLE

Submitted (16-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## COMMENT

On Mar 1, 2002 this sequence version replaced gi:18449677.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT.

FEATURES	Location/Qualifiers
source	1. .163111 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="3" /clone="RP11-301G23"
repeat_region	121. .431 /rpt_family="AluY"
repeat_region	1709. .1765 /rpt_family="MIR"
repeat_region	2495. .3244 /rpt_family="L1ME2"
repeat_region	complement(3663. .3842) /rpt_family="MER5A"
repeat_region	4566. .4612 /rpt_family="(TATATG)n"
repeat_region	complement(5831. .5920) /rpt_family="MIR"
repeat_region	6108. .6141 /rpt_family="AT_rich"
repeat_region	6940. .7252 /rpt_family="AluY"
repeat_region	complement(7727. .7943) /rpt_family="MIR"
repeat_region	8764. .8795

repeat_region	/rpt_family="AT_rich"
repeat_region	8826. .8856 /rpt_family="AT_rich"
repeat_region	complement(9534. .9670) /rpt_family="MIR"
repeat_region	complement(9671. .10122) /rpt_family="MER45B"
repeat_region	10124. .10873 /rpt_family="L1PA4"
repeat_region	complement(10879. .11461) /rpt_family="MER45B"
repeat_region	12515. .12659 /rpt_family="MIR"
repeat_region	12671. .12948 /rpt_family="AluJb"
repeat_region	13292. .13603 /rpt_family="AluY"
repeat_region	13825. .13965 /rpt_family="L1P"
repeat_region	13961. .15866 /rpt_family="L1PA7"
repeat_region	complement(16016. .16132) /rpt_family="FLAM_C"
STS	17586. .17734 /standard_name="6896"
repeat_region	17735. .17762 /rpt_family="(A)n"
repeat_region	19317. .19411 /rpt_family="L2"
repeat_region	20418. .20620 /rpt_family="MER45C"
repeat_region	20906. .21227 /rpt_family="MER45C"
repeat_region	21963. .22111 /rpt_family="MIR"
repeat_region	23721. .23816 /rpt_family="(CCCA)n"
repeat_region	24296. .24443 /rpt_family="MIR"
repeat_region	complement(25826. .25871) /rpt_family="MIR"
repeat_region	complement(25872. .26178) /rpt_family="AluSg1"
repeat_region	complement(26179. .26349) /rpt_family="MIR"
repeat_region	27742. .27968 /rpt_family="MLT1D"
repeat_region	27969. .28051 /rpt_family="(ATGGTG)n"
repeat_region	28052. .28320 /rpt_family="MLT1D"
repeat_region	29165. .29312 /rpt_family="L1MC5"
repeat_region	29437. .30077 /rpt_family="L1MC5"
repeat_region	30866. .30890 /rpt_family="(TTTA)n"
repeat_region	complement(30891. .31171) /rpt_family="AluJb"
repeat_region	31257. .31306 /rpt_family="AT_rich"
repeat_region	31344. .31378 /rpt_family="AT_rich"
repeat_region	complement(32407. .32644)

Query Match 3.3%; Score 55; DB 9; Length 163111;  
Best Local Similarity 100.0%; Pred. No. 1.1e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 TGGATCACCTGAAGTCAGGAGTTCAAGACTAGCCTGGCCACATGGTGAACCT 177

|||||

Db 150432 TGGATCACCTGAAGTCAGGAGTTCAAGACTAGCCTGGCCACATGGTGAAC:CI 150486

## RESULT 31

AC011825/c

## LOCUS

AC011825 Homo sapiens chromosome 18, clone RP11-5316, linear PRI 19-JUL-2002

## DEFINITION

AC011825 Homo sapiens chromosome 18, clone RP11-5316, complete sequence.

## ACCESSION

AC011825.13 GI:21909530

## VERSION

HTG.

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (19-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jul 19, 2002 this sequence version replaced gi:21699439.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2559

Center clone name: 53\_I\_6

-----

## FEATURES

## source

1. .176222  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="18"  
 /map="18"  
 /clone="RP11-5316"  
 /clone\_lib="RPC1-11 Human Male BAC"

## repeat\_region

113..223

## repeat\_region

/rpt\_family="CT-rich"

## repeat\_region

complement(304..597)

## repeat\_region

/rpt\_family="AluYa5"

## repeat\_region

1544..1831

## repeat\_region

/rpt\_family="AluX"

## repeat\_region

complement(2664..2973)

## repeat\_region

/rpt\_family="AluSg"

## repeat\_region

complement(2974..3211)

## repeat\_region

/rpt\_family="MER33"

## repeat\_region

complement(3212..3290)

## repeat\_region

/rpt\_family="AluSp/q"

## repeat\_region

3510..3809

## repeat\_region

/rpt\_family="AluX"

## repeat\_region

3897..3944

## repeat\_region

/rpt\_family="MER33"

## repeat\_region

3946..4250

## repeat\_region

/rpt\_family="AluX"

## repeat\_region

4686..4912

## repeat\_region

/rpt\_family="AluSg/x"

## repeat\_region

4984..5284

## repeat\_region

/rpt\_family="AluY"

## repeat\_region

complement(5557..5703)

## repeat\_region

/rpt\_family="AluJb"

## repeat\_region

complement(5704..6004)

## repeat\_region

/rpt\_family="AluSc"

## repeat\_region

complement(6135..6268)

## repeat\_region

/rpt\_family="L1MC/D"

## repeat\_region

6439..6511

## repeat\_region

/rpt\_family="TAGA)n"

## repeat\_region

6513..6797

## repeat\_region

/rpt\_family="AluX"

## repeat\_region

6798..7121

## repeat\_region

/rpt\_family="AluSg"

## repeat\_region

complement(7052..7105)

## repeat\_region

/note="single clone coverage"

## repeat\_region

complement(7148..7327)

## repeat\_region

/rpt\_family="MIR"

## repeat\_region

8004..8087

## repeat\_region

/rpt\_family="MIR"

## repeat\_region

complement(8993..9305)

## repeat\_region

/rpt\_family="AluJo"

## repeat\_region

9306..9330

## repeat\_region

/rpt\_family="AT-rich"

## repeat\_region

9512..9547

## repeat\_region

/rpt\_family="L2"

## repeat\_region

complement(9548..9686)

## RESULT 31

AC011825/c

## LOCUS

AC011825 Homo sapiens chromosome 18, clone RP11-5316, linear PRI 19-JUL-2002

## DEFINITION

AC011825 Homo sapiens chromosome 18, clone RP11-5316, complete sequence.

## ACCESSION

AC011825.13 GI:21909530

## VERSION

HTG.

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

AC011825 176222 bp DNA linear PRI 19-JUL-2002  
 Homo sapiens chromosome 18, clone RP11-5316, complete sequence.  
 AC011825  
 AC011825.13 GI:21909530  
 HTG.  
 Homo sapiens.  
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 176222)

Birren,B., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 18, clone RP11-5316

Unpublished

2 (bases 1 to 176222)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,

Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,

Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,

Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,

Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,

Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,

Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,

Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,

McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,

Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,J.,

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,

Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,

Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 176222)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,

Barn,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,

Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,

Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,

Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,

Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,

Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,

Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,

McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,

O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,

Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,

Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,

Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,

Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,

Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 176222)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,

Barn,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,

Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,

Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,

Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,

Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,

Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,

Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,

McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,

O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,

Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,

Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,

Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,

Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,

Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 176222)

Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,

Barn,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,

Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,

Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,

Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,

Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,

Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,

Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,

McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,

O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,

Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,

Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,

Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,

Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,

Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.





STS  
dbSTS:STS22742  
Identified using the e-PCR software (G. Schuler)"  
72824..72973  
/note="matching EMBL:F05082  
RHdb:RH100955  
RHdb:RH53917  
RHdb:RH14408  
dbSTS:STS13281  
Identified using the e-PCR software (G. Schuler)"  
73279..73383  
/note="matching EMBL:Z28637  
RHdb:RH9829  
dbSTS:STS4554  
Identified using the e-PCR software (G. Schuler)"  
73374..73474  
/note="matching EMBL:M86162  
RHdb:RH28647  
dbSTS:STS18435  
Identified using the e-PCR software (G. Schuler)"  
74930..75079  
/note="matching EMBL:Z38389  
RHdb:RH44245  
RHdb:RH11243  
dbSTS:STS9878  
Identified using the e-PCR software (G. Schuler)"  
74933..75034  
/note="matching EMBL:Z38389  
RHdb:RH26783  
dbSTS:STS16763  
Identified using the e-PCR software (G. Schuler)"  
75060..75190  
/note="matching EMBL:R26115  
RHdb:RH44163  
RHdb:RH11167  
dbSTS:STS30136  
Identified using the e-PCR software (G. Schuler)"  
75174..75339  
/note="matching EMBL:TL6161  
RHdb:RH53638  
dbSTS:STS20651  
Identified using the e-PCR software (G. Schuler)"  
99195..99349  
/note="matching EMBL:G07510  
RHdb:RH96271  
RHdb:RH31992  
dbSTS:STS10310  
Identified using the e-PCR software (G. Schuler)"  
BASE COUNT 47240 a 40904 c 41487 g 49203 t  
ORIGIN  
Query Match 3.3%; Score 55; DB 9; Length 178834;  
Best Local Similarity 100.0%; Pred. No. 1.1e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 213 GCACACCTGTAAATCCAGCTACTAGGAGCTGAGCAGAGAGAAATTCCTTGAA 267  
|||||  
Db 96181 GCACACCTGTAAATCCAGCTACTAGGAGCTGAGCAGAGAGAAATTCCTTGAA 96127  
RESULT 33  
AC006536/c 185402 bp DNA linear PRI 21-JUL-1999  
LOCUS Homo sapiens chromosome 14 clone BAC257P13 map 14q31, complete  
DEFINITION sequence.  
AC006536  
VERSION AC006536.2 GI:5542034  
KEYWORDS HTG.  
SOURCE HOMO SAPIENS.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 185402)  
AUTHORS Madan,A., Rowen,L., Qin,S., Abbasi,N., Dors,M., Dickhoff,R.,

Harrison,G., James,R., Loretz,C., Lasky,S., Madan,A., Prescott,S., Ratcliffe,A., Shaffer,T. and Hood,L.  
Sequencing of human chromosome 14  
2 (bases 1 to 185402)  
Rowen,L., Madan,A., Qin,S., Abbasi,N., Dors,M., Dickhoff,R., Hicks,P., James,R., Loretz,C., Lasky,S., Madan,A., Ratcliffe,A., Shaffer,T. and Hood,L.  
Direct Submission  
Submitted (07-FEB-1999) Multimegabase Sequencing Center, University of Washington, PO BOX 357730, Seattle, WA 98195, USA  
3 (bases 1 to 185402)  
Madan,A., Rowen,L., Qin,S., Abbasi,N., Dors,M., Dickhoff,R., Harrison,G., James,R., Loretz,C., Lasky,S., Madan,A., Prescott,S., Ratcliffe,A., Shaffer,T. and Hood,L.  
Direct Submission  
Submitted (21-JUL-1999) Multimegabase Sequencing Center, University of Washington, PO BOX 357730, Seattle, WA 98195, USA  
On Jul 21, 1999 this sequence version replaced gi.4235135.  
FEATURES  
Location/Qualifiers  
1..185402  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="14"  
/map="14q31"  
/clone="BAC257P13"  
/clone\_lib="RPCI-11"  
23576..31407  
/gene="calmodulin"  
join(23661..23663,26493..26523,27687..27830,30290..30396,30807..30942,31117..31145)  
/gene="calmodulin"  
/note="Intron-exon boundaries defined in relation to cDNA in M27319."  
/codon\_start=1  
/product="calmodulin"  
/protein\_id="AAD45181.1"  
/db\_xref="GI:5542035"  
TEAELODMINEVDADNGTIDFPFELIMARKMKADTSEEREAFRVFKDNGNYIS  
AAELRHVMTNLGERLTDDEVDREADIDGQVNYEEFVQMMTK"  
/translation="MADQLTEEOIAEKEAFSLFDKDGDTITTKELGTVMRSLGQNP  
107150..107270  
/note="low quality data"  
122750..123220  
/note="low quality data -- this region was resistant to several finishing strategies, especially in the microsatellite"  
132730..132810  
/note="low quality data"  
134050..134150  
/note="low quality data"  
144350..144550  
/note="low quality data"  
165170..165260  
/note="low quality data"  
BASE COUNT 46161 a 45067 c 45138 g 49036 t  
ORIGIN  
Query Match 3.3%; Score 55; DB 9; Length 185402;  
Best Local Similarity 100.0%; Pred. No. 1.1e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 213 GCACACCTGTAAATCCAGCTACTAGGAGCTGAGCAGAGAGAAATTCCTTGAA 267  
|||||  
Db 55530 GCACACCTGTAAATCCAGCTACTAGGAGCTGAGCAGAGAGAAATTCCTTGAA 55476  
RESULT 34  
AC015563/c 214984 bp DNA linear PRI 31-JUL-2002  
LOCUS Homo sapiens chromosome 18, clone RP11-344B2, complete sequence.  
DEFINITION  
AC015563  
ACCESSION  
VERSION  
AUTHORS AC015563.11 GI:220245600

KEYWORDS  
SOURCE  
ORGANISM

HTG.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 214984)  
Birren,B., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 18, clone RP11-344B2

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

2 (bases 1 to 214984)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,  
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Doneelan,L., Doyle,M.,  
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,  
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wymao,D., Ye,W.J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 214984)  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (06-JUL-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 214984)  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (31-JUL-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 31, 2002 this sequence version replaced gi:21699684.  
All repeats were identified using RepeatMasker:

FEATURES  
source

Location/Qualifiers  
1. 214984  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18"  
/clone="RP11-344B2"  
/clone\_lib="RPC1-11 Human Male BAC"  
complement(683..970)  
/rpt\_family="AluSx"  
complement(980..1266)  
/rpt\_family="AluSx"  
1595..1656  
/rpt\_family="MIR"  
complement(1823..1911)  
/rpt\_family="AluJb"  
complement(1904..1908)  
/note="<30 qual SNGL region"  
complement(1925..2041)  
/note="<30 qual SNGL region"  
complement(1935..2003)  
/rpt\_family="FRAM/FAM"  
complement(1960)  
/note="probably G, possibly C"  
complement(2070)  
/note="probably C, possibly A"  
complement(2073..2079)  
/note="<30 qual SNGL region"  
complement(2084..2088)  
/note="<30 qual SNGL region"  
complement(2097..2101)  
/note="<30 qual SNGL region"  
2109..2184  
/rpt\_family="AluSp/q"  
complement(2111..2141)  
/note="<30 qual SNGL region"  
complement(2574..2653)  
/rpt\_family="Charliel"  
2654..3007  
/rpt\_family="MER7A"  
complement(3008..3047)  
/rpt\_family="Charliel"  
3048..3181  
/rpt\_family="FLAM\_C"  
complement(3182..3477)  
/rpt\_family="Charliel"  
3478..3781  
/rpt\_family="AluJb"  
complement(3782..3832)  
/rpt\_family="Charliel"  
complement(3835..3974)  
/rpt\_family="AluSp/q"  
complement(3975..4559)  
/rpt\_family="Charliel"  
complement(4566..4721)  
/rpt\_family="AluSg/x"  
complement(4743..4933)  
/rpt\_family="L1MC/D"  
4994..5147  
/rpt\_family="L1MC/D"  
complement(5227..5616)  
/rpt\_family="L1MC/D"

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT



RHdb:RH31992  
dbSTS:STS10310  
Identified using the e-PCR software (G. Schuler)"

BASE COUNT 58286 a 55369 c 56129 g 58313 t

ORIGIN

Query Match 3.3%; Score 55; DB 9; Length 228097;  
Best Local Similarity 100.0%; Pred. No. 1e-17;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 GCACACCTGTAAATCCAGCTACTTAGGAGCTGAGCGAGGAGATTGCTTCAA 267  
|||||  
Db 220169 GCACACCTGTAAATCCAGCTACTTAGGAGCTGAGCGAGGAGATTGCTTCAA 220223

RESULT 36  
AF228730

LOCUS AF228730 265308 bp DNA linear PRI 25-JUL-2002

DEFINITION Homo sapiens chromosome 8 clone CTB-415D8, SCB-332a23 map 8p22-p21, complete sequences.

ACCESSION AF228730 AF231129

VERSION AF228730.7 GI:21954989

KEYWORDS HTG.

SOURCE Homo sapiens.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 265308)  
Schudy,A., Schilhabel,M., Baumgart,C., Menzel,U., Weber,J.,  
Schattevoy,R. and Rosenthal,A.  
Unpublished  
2 (bases 1 to 265308)  
Reichwald,K. and Platzer,M.  
Direct Submission  
Submitted (23-AUG-2001) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany  
3 (bases 1 to 265308)  
Siddiqui,R. and Platzer,M.  
Direct Submission  
Submitted (05-APR-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
4 (bases 1 to 265308)  
Heinze,I. and Platzer,M.  
Direct Submission  
Submitted (28-JUN-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
5 (bases 1 to 265308)  
Lagemann,D. and Platzer,M.  
Direct Submission  
Submitted (25-JUL-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany  
On Jul 25, 2002 this sequence version replaced gi:21622700.

----- Genome Center  
Center: Institute of Molecular Biotechnology  
Center code: IMB  
Web site: <http://genome.imb-jena.de/>  
Contact: gscj-submit@genome.imb-jena.de  
----- Project Information  
Center project name: H150+H369  
Center clone name: CTB-415D8, SCB-332a23  
----- Summary Statistics  
Sequencing vector: pUC18; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 264684 bases at least Q40  
Consensus quality: 265117 bases at least Q30  
Consensus quality: 265237 bases at least Q20  
Quality coverage: 12.19x  
-----  
This sequence was finished as follows unless otherwise noted: all  
regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest.

-----  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
Genbank flat file format but are available as part  
of this entry's ASN.1 file.

-----  
Location/Qualifiers  
1. 265308  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8p22-p21, complete sequences"  
/clone="CTB-415D8, SCB-332a23"  
1. 200903  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
96358. 265308  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="SCB-332a23"  
6848. 6901  
/note="single stranded/single chemistry region"  
8852. 8891  
/note="single stranded/single chemistry region"  
14189. 14407  
/note="single stranded/single chemistry region"  
16957. 17120  
/note="single stranded/single chemistry region"  
21751. 22252  
/note="single stranded/single chemistry region"  
22253. 22295  
/note="single stranded/single chemistry region"  
22788. 22922  
/note="pcr product sequence only , CTB-415D8"  
23329. 23373  
/note="single stranded/single chemistry region"  
31322. 31787  
/note="single stranded/single chemistry region"  
35136. 35258  
/note="single stranded/single chemistry region"  
36846. 37017  
/note="single stranded/single chemistry region"  
40133. 40293  
/note="single stranded/single chemistry region"  
52565. 52706  
/note="single stranded/single chemistry region"  
59224. 59480  
/note="single stranded/single chemistry region"  
59605. 59892  
/note="single stranded/single chemistry region"  
60332. 60551  
/note="single stranded/single chemistry region"  
61112. 61470  
/note="single stranded/single chemistry region"  
63246. 63466  
/note="single stranded/single chemistry region"  
63433. 63434  
/note="low quality region , CTB-415D8"  
63467. 63529  
/note="single clone coverage , low quality region ,  
CTB-415D8"  
63530. 63541  
/note="single stranded/single chemistry region"  
66729. 66731  
misc\_feature

misc\_feature /note="single clone coverage"  
66732 /note="single clone coverage , low quality region ,  
CTB-415D8"  
misc\_feature 66733..66737 /note="single clone coverage"  
misc\_feature 66847..66925 /note="single stranded/single chemistry region"  
misc\_feature 66921..66925 /note="low quality region , CTB-415D8"  
misc\_feature 66957..66958 /note="low quality region , CTB-415D8"  
misc\_feature 67550..67655 /note="single stranded/single chemistry region"  
misc\_feature 67579..67580 /note="low quality region , CTB-415D8"  
misc\_feature 67584 /note="low quality region , CTB-415D8"  
misc\_feature 67656..67666 /note="single stranded/single chemistry region"  
misc\_feature 67667..67697 /note="single clone coverage , low quality region ,  
CTB-415D8"  
misc\_feature 67698..67744 /note="low quality region , CTB-415D8"  
misc\_feature 67698..67730 /note="single stranded/single chemistry region"  
misc\_feature 67731..67819 /note="single stranded/single chemistry region"  
misc\_feature 67820..67825 /note="single stranded/single chemistry region"  
misc\_feature 69072..69252 /note="single stranded/single chemistry region"  
misc\_feature 73056..73722 /note="single stranded/single chemistry region"  
misc\_feature 73723..74005 /note="single stranded/single chemistry region"  
misc\_feature 74006..74061 /note="single stranded/single chemistry region"  
misc\_feature 78932 /note="low quality region , CTB-415D8"  
misc\_feature 80878..81170 /note="single stranded/single chemistry region"  
misc\_feature 81171..81249 /note="single stranded/single chemistry region"  
misc\_feature 83988..84084 /note="single stranded/single chemistry region"  
misc\_feature 86758..86824 /note="single stranded/single chemistry region"  
misc\_feature 86769 /note="low quality region , CTB-415D8"  
misc\_feature 87333 /note="low quality region , CTB-415D8"  
misc\_feature 89180..89507 /note="single stranded/single chemistry region"  
misc\_feature 90464 /note="low quality region , CTB-415D8"  
misc\_feature 90465 /note="low quality region , CTB-415D8"  
misc\_feature 90467..90803 /note="single stranded/single chemistry region"  
misc\_feature 90700 /note="low quality region , CTB-415D8"  
misc\_feature 90701 /note="low quality region , CTB-415D8"  
misc\_feature 90734 /note="low quality region , CTB-415D8"  
misc\_feature 90741 /note="low quality region , CTB-415D8"  
misc\_feature 90742..90745 /note="low quality region , CTB-415D8"  
Query Match 3.2%; Score 54; DB 9; Length 265308;

Best Local Similarity 100.0%; Pred. No. 3.3e-17;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 206 CTGTTGGGCACACCTGTAAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAAT 259  
|||||  
Db 143104 CGTGGTGGGCACACCTGTAAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAAT 143157  
|||||  
RESULT 37  
AC025275 AC025275 94850 bp DNA linear PRI 01-DEC-2000  
LOCUS Homo sapiens chromosome 16 clone CTC-502C3, complete sequence.  
DEFINITION AC025275  
ACCESSION AC025275  
VERSION AC025275.4 GI:11496340  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 94850)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 94850)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (08-MAR-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 94850)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (01-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Dec 1, 2000 this sequence version replaced gi:8576144.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.8% of Sequence;  
Estimated Total Number of Errors is 0.2.  
STS Content:  
SHGC-60502 G36841.  
Location/Qualifiers  
1..94850  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="CTC-502C3"  
BASE COUNT 23899 a 22519 c 22946 g 25486 t  
ORIGIN  
Query Match 3.2%; Score 53; DB 9; Length 94850;  
Best Local Similarity 100.0%; Pred. No. 1.3e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 207 CTGTTGGGCACACCTGTAAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAAT 259  
|||||  
Db 9909 GTGTTGGGCACACCTGTAAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAAT 9961  
|||||  
RESULT 38  
HUAC002551/C HUAC002551 114411 bp DNA linear PRI 23-NOV-1999  
LOCUS Human Chromosome 16 BAC clone CIT987SK-A-951C11, complete sequence.  
DEFINITION AC002551  
ACCESSION AC002551  
VERSION AC002551.1 GI:2809276  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 114411)  
AUTHORS Loftus,B.J., Kim,U.J., Sneddon,V.P., Kalush,F., Brandon,R.,

Fuhrmann, J., Mason, T., Crosby, M.L., Barnstead, M., Cronin, L., Deslattes Mays, A., Cao, Y., Xu, R.X., Kang, H.L., Mitchell, S., Eichler, E., Harris, P.C., Venter, J.C. and Adams, M.D.  
Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q  
Genomics 60 (3), 295-308 (1999)  
99425270  
10493829  
2 (bases 1 to 114411)  
Adams, M.D., Loftus, B.J., Zhou, L., Crosby, M., Fuhrmann, J., Brandon, R., Kim, U.J., Kerlavage, A.R. and Venter, J.C.  
Human Chromosome 16 BAC clone CIT987SK-A-951C11  
Unpublished  
3 (bases 1 to 114411)  
Adams, M.D.  
Direct Submission  
Submitted (19-SEP-1997) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
4 (bases 1 to 114411)  
Adams, M.D.  
Direct Submission  
Submitted (26-JAN-1998) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
5 (bases 1 to 114411)  
Adams, M.D.  
Direct Submission  
Submitted (28-JAN-1998) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
6 (bases 1 to 114411)  
Adams, M.D. and Loftus, B.J.  
Direct Submission  
Submitted (17-MAR-1998) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA  
On Jan 26, 1998 this sequence version replaced gi:2642422.  
Address all correspondence to: Mark Adams The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA e-mail address: mdadams@tigr.org. The orientation of the sequence is from SP6 end to T7 end. Genes were identified by a combination of five methods including: XGRAIL (available by anonymous ftp from arthur.epm.ornl.gov), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, http://gnomic.stanford.edu/~chris/GNSCANW.html) searches of the complete sequence against a peptide database and the Human gene Index database at TIGR (http://www.tigr.org/tdb/hg1/hgi.html). Genes without peptide homology having spliced EST hits are termed 'Unknown gene product'. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, http://genome.wustl.edu/eddy/tRNAscan-SE/).

## FEATURES

Location/Qualifiers  
1..114411  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/map="16p12.1"  
/clone="A-951C11"  
/complement(<3210..3493)  
/gene="A-951C1A1."  
complement(<3210..3493)  
/gene="A-951C1A1."  
complement(<3210..3430)  
/gene="A-951C1A1."  
/codon\_start=1  
/products="transcription factor (TFIIIC) alpha chain, (3' partial)"  
/protein\_id="AAC05811.1"  
/db\_xref="GI:2833636"  
/translation="MDALFSLDVALEGLDGLPALMSRLTRVPPFPLPEPCTQ  
EFLWALATHPGISFYEPFRPDLQLQDR"  
6327..6444

## STS

STS  
/db\_xref="dbSTS:G02122"  
101138..101255  
/db\_xref="dbSTS:G02122"  
BASE COUNT 29730 a 26699 c 27271 g 30711 t  
ORIGIN  
Query Match 3.2%; Score 53; DB 9; Length 114411;  
Best Local Similarity 100.0%; Pred. No. 1.3e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 207 GTGGTGGCACACACCTGTAAATCCAGCTACTAGGAGGCTGAGCAGGAGAAT 259  
|||||  
Db 11842 GTGGTGGCACACACCTGTAAATCCAGCTACTAGGAGGCTGAGCAGGAGAAT 11790  
RESULT 39  
AC034128 149628 bp DNA linear HTG 04-MAY-2000  
LOCUS Homo sapiens chromosome 8 clone RP11-207E12 map 8, WORKING DRAFT  
DEFINITION AC034128  
AC034128  
VERSION AC034128.2 GI:7705189  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 149628)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 8, clone RP11-207E12  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 149628)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F., Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArefilano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, D., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grant-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Meneses, L., Mihova, T., Miranda, C., Miñana, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Thodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, W.  
Direct Submission  
Submitted (04-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 4, 2000 this sequence version replaced gi:7408059.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www.seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L9020  
Center clone name: 207\_E12  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 141708 bases at least Q40

Consensus quality: 145382 bases at least Q30  
Consensus quality: 147046 bases at least Q20  
Insert size: 154000; agarose-fp  
Quality coverage: 4.8 in Q20 bases; agarose-fp  
Quality coverage: 5.0 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1101: contig of 1101 bp in length  
1102 1201: gap of 100 bp  
1202 2603: contig of 1402 bp in length  
2604 2703: gap of 100 bp  
2704 4989: contig of 2286 bp in length  
4990 5089: gap of 100 bp  
5090 6924: contig of 1835 bp in length  
6925 7024: gap of 100 bp  
7025 10629: contig of 3605 bp in length  
10630 10729: gap of 100 bp  
10730 14318: contig of 3589 bp in length  
14319 14418: gap of 100 bp  
14419 18825: contig of 4407 bp in length  
18826 18925: gap of 100 bp  
18926 22896: contig of 3971 bp in length  
22897 22996: gap of 100 bp  
22997 26636: contig of 3640 bp in length  
26637 26736: gap of 100 bp  
26737 31014: contig of 4278 bp in length  
31015 31114: gap of 100 bp  
31115 35710: contig of 4596 bp in length  
35711 35810: gap of 100 bp  
35811 41703: contig of 5893 bp in length  
41704 41803: gap of 100 bp  
41804 55435: contig of 13632 bp in length  
55436 55535: gap of 100 bp  
55536 73495: contig of 17960 bp in length  
73496 73595: gap of 100 bp  
73596 95124: contig of 21529 bp in length  
95125 95224: gap of 100 bp  
95225 115620: contig of 20396 bp in length  
115621 115720: gap of 100 bp  
115721 149628: contig of 33908 bp in length.

FEATURES  
Source

Location/Qualifiers  
1. .149628  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-207E12"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. .1101  
/note="assembly\_fragment"  
1202. .2603  
/note="assembly\_fragment"  
2704. .4989  
/note="assembly\_fragment"  
5090. .6924  
/note="assembly\_fragment"  
7025. .10629  
/note="assembly\_fragment"  
10730. .14318  
/note="assembly\_fragment"  
14419. .18825  
/note="assembly\_fragment"  
18926. .22896  
/note="assembly\_fragment"  
22997. .26636

misc\_feature /note="assembly\_fragment"  
26737. .31014  
misc\_feature /note="assembly\_fragment"  
31115. .35710  
misc\_feature /note="assembly\_fragment"  
35811. .41703  
misc\_feature /note="assembly\_fragment"  
41804. .55435  
misc\_feature /note="assembly\_fragment"  
55536. .73495  
misc\_feature /note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right"  
73596. .95124  
misc\_feature /note="assembly\_fragment"  
95225. .115620  
misc\_feature /note="assembly\_fragment"  
115721. .149628  
misc\_feature /note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right"  
BASE COUNT 47928 a 26944 c 26014 g 47142 t 16000 others  
ORIGIN  
Query Match 3.2%; Score 53; DB 2; Length 149628;  
Best Local Similarity 100.0%; Pred. No. 1.2e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 215 ACACACCTGTAATCCAGGCTACTTAGGAGGCTAGGAGGAGAGAAATTCGTTGAA 267  
|||||  
Db 136014 ACACACCTGTAATCCAGGCTACTTAGGAGGCTAGGAGGAGAGAAATTCGTTGAA 136066  
RESULT 40  
AL391683  
LOCUS  
DEFINITION Human DNA sequence from clone Rp11-184I23 on chromosome 13 Contains  
STSS and GSSs, complete sequence.  
ACCESSION AL391683  
VERSION AL391683.8 GI:11022494  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Dunn,M.  
REFERENCE 1 (bases 1 to 168373)  
AUTHORS  
TITLE Direct Submission  
JOURNAL Submitted (16-MAR-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk  
COMMENT On Oct 25, 2000 this sequence version replaced gi:10933277.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C-elegans/wormpep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 13, constructed by the Sanger Centre Chromosome 13  
Mapping Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chrl3  
This sequence is the entire insert of clone Rp11-184I23 The true  
left end of clone Rp11-94N9 is at 82505 in this sequence. This  
sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such



as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-184I23 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6

#### FEATURES

##### source

location/Qualifiers  
1..168373  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="13"  
/clone="RP11-184I23"  
/clone\_lib="RPCI-11.1"

repeat\_region  
1..2005  
/note="L1PA15 repeat: matches 4126. .6153 of consensus"  
repeat\_region  
2452..2804  
/note="MLT1A1 repeat: matches 12. .365 of consensus"  
repeat\_region  
2814..2924  
/note="L2 repeat: matches 2633. .2750 of consensus"  
repeat\_region  
3367..3667  
/note="AluY repeat: matches 1. .303 of consensus"  
misc\_feature  
complement(3885..4359)  
/note="match: GSS: Em:AQ834807"  
misc\_feature  
4360..4937  
/note="match: GSS: Em:AQ503925"  
repeat\_region  
5453..5926  
/note="MLT1D repeat: matches 2. .505 of consensus"  
repeat\_region  
7710..8338  
/note="HAL1 repeat: matches 918. .1591 of consensus"  
repeat\_region  
8339..8452  
/note="57 copies 2 mer at 73% conserved"  
repeat\_region  
8602..8920  
/note="HAL1 repeat: matches 407. .747 of consensus"  
repeat\_region  
9060..9498  
/note="Charlieb repeat: matches 51. .522 of consensus"  
repeat\_region  
9634..9763  
/note="HAL1 repeat: matches 31. .162 of consensus"  
repeat\_region  
10404..10489  
/note="43 copies 2 mer ta 74% conserved"  
repeat\_region  
12964..13333  
/note="WSTD repeat: matches 1. .387 of consensus"  
repeat\_region  
13504..13648  
/note="MER44C repeat: matches 1. .147 of consensus"  
repeat\_region  
13649..13823  
/note="FRAM repeat: matches 2. .159 of consensus"  
repeat\_region  
13824..14411  
/note="MER44C repeat: matches 147. .723 of consensus"  
repeat\_region  
14891..15247  
/note="MLT1A1 repeat: matches 1. .359 of consensus"  
repeat\_region  
15999..16636  
/note="L2 repeat: matches 710. .1436 of consensus"  
repeat\_region  
16936..17037  
/note="L2 repeat: matches 1808. .1912 of consensus"  
repeat\_region  
17581..17604  
/note="12 copies 2 mer tt 100% conserved"  
repeat\_region  
18266..18713  
/note="LTR40b repeat: matches 4. .459 of consensus"  
misc\_feature  
complement(18746..19209)  
/note="match: GSS: Em:AQ053565"  
misc\_feature  
18779..19249  
/note="match: GSS: Em:AQ469501"  
repeat\_region  
18834..18973  
/note="MER41B repeat: matches 157. .309 of consensus"  
misc\_feature  
complement(18932..19307)  
/note="match: GSS: Em:AQ343749"  
repeat\_region  
19716..19897  
/note="L1MC5 repeat: matches 7739. .7927 of consensus"  
repeat\_region  
20174..20221  
/note="24 copies 2 mer tt 77% conserved"  
repeat\_region  
21236..21484  
/note="L1P83 repeat: matches 5922. .6150 of consensus"  
repeat\_region  
22398..22480

repeat\_region  
22557..22757  
/note="MIR repeat: matches 25. .221 of consensus"  
repeat\_region  
24392..24478  
/note="MIR repeat: matches 169. .256 of consensus"  
repeat\_region  
24875..25007  
/note="MER57-internal repeat: matches 7152. .7288 of consensus"  
repeat\_region  
25565..25643  
/note="L1P3 repeat: matches 607. .685 of consensus"  
repeat\_region  
25666..26287  
/note="LTR12 repeat: matches 17. .538 of consensus"  
repeat\_region  
26294..28555  
/note="L1PA12 repeat: matches 764. .1499 of consensus"  
repeat\_region  
28556..28682  
/note="L1PA7 repeat: matches 6017. .6145 of consensus"  
repeat\_region  
28683..31036  
/note="L1PA12 repeat: matches 1499. .5199 of consensus"  
repeat\_region  
31038..31484  
/note="HERV repeat: matches 3942. .4388 of consensus"  
repeat\_region  
31485..31795  
/note="AluY repeat: matches 1. .309 of consensus"  
repeat\_region  
31796..31817  
/note="HERV repeat: matches 3921. .3942 of consensus"  
repeat\_region  
32005..32198  
/note="HERVE repeat: matches 2138. .2321 of consensus"  
repeat\_region  
32547..33187  
/note="HERVE repeat: matches 1217. .1855 of consensus"  
repeat\_region  
33615..33649  
/note="HERVE repeat: matches 808. .842 of consensus"  
repeat\_region  
34416..34625  
/note="HSR1RT repeat: matches 1. .202 of consensus"  
repeat\_region  
34754..35029  
/note="LTR30 repeat: matches 235. .529 of consensus"  
misc\_feature  
36368..36949  
/note="match: GSS: Em:AQ309823"  
repeat\_region  
37296..37407  
/note="L2 repeat: matches 1194. .1311 of consensus"  
repeat\_region  
37958..38196  
/note="MIR repeat: matches 15. .262 of consensus"  
repeat\_region  
38580..38880  
/note="AluY repeat: matches 1. .302 of consensus"  
repeat\_region  
39097..39178  
/note="MER81 repeat: matches 5. .83 of consensus"  
repeat\_region  
41431..41928  
/note="L2 repeat: matches 2237. .2750 of consensus"  
repeat\_region  
41933..42381  
/note="MLT1C repeat: matches 10. .461 of consensus"  
repeat\_region  
42387..42923  
/note="L2 repeat: matches 1733. .2239 of consensus"  
misc\_feature  
44677..45206  
/note="match: GSS: Em:AQ384058"  
repeat\_region  
44790..44894  
/note="MIR repeat: matches 28. .138 of consensus"  
repeat\_region  
45703..46000  
/note="L2 repeat: matches 2137. .2419 of consensus"  
misc\_feature  
complement(46088..46407)  
/note="match: GSS: Em:AQ631086"  
misc\_feature  
complement(46093..46316)  
/note="match: GSS: Em:A2098853"  
misc\_feature  
complement(46093..46273)  
/note="match: GSS: Em:A2060650"  
misc\_feature  
complement(46113..46453)  
/note="match: GSS: Em:AQ061481"  
misc\_feature  
46114..46438  
/note="match: GSS: Em:AQ598220"  
misc\_feature  
complement(46114..46266)  
/note="match: GSS: Em:AQ388769"  
misc\_feature  
46126..46325  
/note="match: GSS: Em:A2308449"  
misc\_feature  
46154..46390  
/note="match: GSS: Em:AQ321139"  
misc\_feature  
complement(46176..46359)

```

/note="match: STS: Em:HSA246YG9"
complement(46176..46316)
/note="match: GSS: Em:AQ807289"
46208..46450
/note="match: GSS: Em:AQ017857"
46211..46272
/note="31 copies 2 mer at 77% conserved"
complement(46218..46453)
/note="match: GSS: Em:AQ822390"
complement(join(46222..46326,54553..54701))
/note="match: GSS: Em:AZ030834"
complement(46222..46460)
/note="match: GSS: Em:AQ424188"
complement(46222..46453)
/note="match: GSS: Em:AQ628671"
complement(46229..46412)

Query Match      3.2%; Score 53; DB 9; Length 168373;
Best Local Similarity 100.0%; Pred. No. 1.2e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 GTGGTGACACACCTCTAATCCAGCTACTTAGGAGGCTGAGCAGGAGAAAT 259
|||||
Db 118575 GTGGTGACACACCTCTAATCCAGCTACTTAGGAGGCTGAGCAGGAGAAAT 118627

RESULT 41
AC025821      178378 bp      DNA      linear      PRI 09-JAN-2002
LOCUS      Homo sapiens BAC clone RP11-472F4 from 4, complete sequence.
DEFINITION      AC025821
ACCESSION      HTG.
VERSION      AC025821.7 GI:15778778
KEYWORDS
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 178378)
Sulston,J.E. and Waterston,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
MEDLINE
PUBMED
9847074
2 (bases 1 to 178378)
Nguyen,C., Haakenson,W. and Phillips,A.
The sequence of Homo sapiens BAC clone RP11-472F4
Unpublished (2001)
3 (bases 1 to 178378)
Waterston,R.H.
Direct Submission
Submitted (15-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 178378)
Waterston,R.H.
Direct Submission
Submitted (26-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 178378)
Waterston,R.
Direct Submission
Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 26, 2001 this sequence version replaced gi:15528932.
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
-----
Summary Statistics
-----
Center project name: H_NH0472F04
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tatenio,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-485C11; the clone sequenced to the right is RP11-319N12. Actual start of this clone is at base position 1 of RP11-472F4; actual end is at base position 178378 of RP11-472F4.

Data from AC023246 and AC019300 was used to finish this clone, AC025821. Polymorphisms have been identified between AC023246 (H\_NH0485C11), AC019300 (H\_NH0319N12), and AC025821 (H\_NH0472F04).

FEATURES	source
repeat_region	1..178378
repeat_region	/organism="Homo sapiens"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="3"
repeat_region	/map="4"
repeat_region	/clone="RP11-472F4"
repeat_region	/clone_lib="RPCI-11"
repeat_region	389..416
repeat_region	/rpt_family="(T)n"
repeat_region	1007..1033
repeat_region	/rpt_family="AT-rich"
repeat_region	1643..1675
repeat_region	/rpt_family="(TTTA)n"
repeat_region	1648..1953
repeat_region	/rpt_family="Alu"
repeat_region	2725..2848
repeat_region	/rpt_family="L1"
repeat_region	3091..3130
repeat_region	/rpt_family="AT-rich"
repeat_region	3424..3457
repeat_region	/rpt_family="AT-rich"
repeat_region	3657..3679
repeat_region	/rpt_family="AT-rich"
repeat_region	5262..5293
repeat_region	/rpt_family="(CA)n"
repeat_region	5975..6059
repeat_region	/rpt_family="GA-rich"
repeat_region	7543..7572
repeat_region	/rpt_family="AT-rich"
repeat_region	8272..8314
repeat_region	/rpt_family="MER53"
repeat_region	9396..9522
repeat_region	/rpt_family="(TA)n"



\* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1
* 7010 7109: contig of 7009 bp in length
* 7110 7109: gap of unknown length
* 23344 23444: contig of 16235 bp in length
* 23345 23444: gap of unknown length
* 23445 38890: contig of 15446 bp in length
* 38891 38990: gap of unknown length
* 38991 55546: contig of 16556 bp in length
* 55547 55647: gap of unknown length
* 55647 103534: contig of 47888 bp in length
* 103535 103635: gap of unknown length
* 103635 181438: contig of 77804 bp in length.
  Location/Qualifiers
    1..181438
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /chromosome="12"
      /clone="RP11-46H11"
      /clone_lib="RPCI human BAC library 11"
    1..7009
      /note="assembly_name:Contig16"
    7110..23344
      /note="assembly_name:Contig17"
    23445..38890
      /note="assembly_name:Contig18"
    38991..55546
      /note="assembly_name:Contig19"
    55647..103534
      /note="assembly_name:Contig20"
      /clone_end:SP6"
    103635..181438
      /note="assembly_name:Contig21"
      /clone_end:17"
  
```

## FEATURES

source

```

misc_feature
  /note="assembly_name:Contig16"
misc_feature
  /note="assembly_name:Contig17"
misc_feature
  /note="assembly_name:Contig18"
misc_feature
  /note="assembly_name:Contig19"
misc_feature
  /note="assembly_name:Contig20"
  /clone_end:SP6"
misc_feature
  /note="assembly_name:Contig21"
  /clone_end:17"
  
```

```

BASE COUNT 44022 a 44380 c 46244 g 46284 t 508 others
ORIGIN
  
```

```

Query Match          3.2%; Score 53; DB 2: Length 181438;
Best Local Similarity 100.0%; Pred No. 1.2e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 GATCACCCTGAAGTCAGGAGTTCAGACTAGCCTGGCCAAACATGCTGAAACCCCT 177
      |||||||
Db 50897 GATCACCCTGAAGTCAGGAGTTCAGACTAGCCTGGCCAAACATGCTGAAACCCCT 50949
      |||||||
  
```

## RESULT 43

```

AC127070
LOCUS          Homo sapiens clone RP11-46H11, *** SEQUENCING IN PROGRESS ***, 4
DEFINITION    AC127070 182908 bp DNA linear HTG 09-AUG-2002
              unordered pieces.
ACCESSION     AC127070.6 GI:22164868
VERSION       HTG; HTGS_PHASE1; HTGS_ACTIVEFIN.
KEYWORDS      human.
SOURCE        Homo sapiens
  
```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  
```

```

1 (bases 1 to 182908)
  
```

```

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbiera,J., Benton,J., Binaige,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,K., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
  
```

```

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudah,S.,
Karlsom,E., Kelly,S., Khan,O., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogih,M., Okwuonu,G.,
Oraguine,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
  
```

## COMMENT

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HEGD
Center clone name: RP11-46H11
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 180515 bases at least Q40
Consensus quality: 180781 bases at least Q30
Consensus quality: 180936 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
      (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 14582: contig of 14582 bp in length
* 14583 14682: gap of unknown length
* 14683 38972: contig of 24290 bp in length
* 38973 39072: gap of unknown length
  
```



```

* 99663 99762: gap of unknown length
* 99763 117806: contig of 18044 bp in length
* 117807 117906: gap of unknown length
* 117907 135151: contig of 17245 bp in length
* 135152 135251: gap of unknown length
* 135252 154753: contig of 19502 bp in length
* 154754 154853: gap of unknown length
* 154854 183456: contig of 28603 bp in length
* 183457 183556: gap of unknown length
* 183557 221574: contig of 38018 bp in length
* 221575 221674: gap of unknown length
* 221675 303011: contig of 81340 bp in length.
FEATURES
    source
    Location/Qualifiers
    1. .303014
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="3"
        /clone="RP11-264F20"
BASE COUNT 94449 a 55614 c 54945 g 95888 t 2118 others
ORIGIN
Query Match 3.2% Score 53; DB 2: Length 303014;
Best Local Similarity 100.0%; Pred. No. 1.1e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 207 GTGTGGCACACCTCTAATCCAGCTACTTAGGAGGCTGAGCGAGGAGAAAT 259
|||||
Db 60597 GTGTGGCACACACCTGTAAATCCAGCTACTTAGGAGGCTGAGCGAGGAGAAAT 60545
|||||

RESULT 45
AP000006/c
LOCUS
DEFINITION Homo sapiens genomic DNA, chromosome 8p11.2, senescence gene
region, section 2/19.
ACCESSION AP000006
VERSION AP000006.1
KEYWORDS AP000006.1 GI:4579987
SOURCE
    ORGANISM
        Homo sapiens DNA.
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1
        Isomura,M., Ikegawa,S., Kinjo,T. and Nakamura,Y.
        DNA sequence analysis of a 1.9-Mb region on chromosome 8p11.2
        Published Only in Database (1999)
REFERENCE
    2 (bases 1 to 100000)
        Hirakawa,M., Yamaguchi,H., Imai,K. and Shimada,J.
        Submitted
        Submitted (12-FEB-1999) Mika Hirakawa, Japan Science and Technology
        Corporation (JST), Advanced Database Department; 5-3, Yonban-cho,
        Chiyoda-ku, Tokyo 102-0028, Japan (E-mail:mika@tokyo.jst.go.jp,
        Tel:81-3-5214-8491, Fax:81-3-5214-8470)
        This sequence is conducted by Japanese Foundation for Cancer
        Research as a JST sequencing team.
        Principal Investigator: Yusuke Nakamura Ph.D
        Phone:81-3-5449-5372, Fax:81-3-5449-5433,
        Yusuke@ngc.ims.u-tokyo.ac.jp
        The sequence is submitted by Human Genome Sequencing in ALIS
        project of JST
        Japan Science and Technology Corporation (JST)
        5-3, Yonbancyo, Chiyoda-ku, Tokyo, 102-0081 Japan
        For further information about this sequences, please visit our
        sequence archive web site (http://www.alis.tokyo.jst.go.jp/HGS/top.
        html) or send email to webmaster@www.alis.tokyo.jst.go.jp.
FEATURES
    source
    Location/Qualifiers
    1. .100000
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="8"
        /map="8p11.2"
        8973. .9159
        /note="SHGC-2770:The location is between each flanking
STS

```

```

site of PCR primers."
/db_xref="GDB:1234502"
BASE COUNT 25261 a 22838 c 23233 g 28644 t 24 others
ORIGIN
Query Match 3.1% Score 52; DB 9: Length 100000;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 208 TGGTGGCACACCTGTAATCCAGCTACTTAGGAGGCTGAGCGAGGAGAAAT 259
|||||
Db 40052 TGGTGGCACACACCTGTAATCCAGCTACTTAGGAGGCTGAGCGAGGAGAAAT 40001
|||||

RESULT 46
AC006208/c
LOCUS
DEFINITION Homo sapiens 3p21.1-9 PAC RPC14-793P23 (Roswell Park Cancer
Institute Human PAC Library) complete sequence.
ACCESSION AC006208
VERSION AC006208.3
KEYWORDS GI:4558540
SOURCE
    ORGANISM
        Homo sapiens.
        Homo sapiens
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
    1 (bases 1 to 123943)
        Muzny,D., Aronson,A.D., Bouck,J., Bunac,C., Chen,J., Chen,Z.,
        Culpepper,P., Ding,Y., Dugan,S.P., Durbin,K.J., Forcum,J.,
        Ganesh,R.P., Garcia,C., Garcia,D.K., Gorrell,H., Gorrell,L.L.,
        He,X., Hernandez,J., Jackson,L.E., Kondejewski,N., Leal,B.,
        Lichtarge,O., Liu,W., Logan,O., Lu,J., Martinez,C., Moore,S.,
        Moorish,T., Nguyen,N., Oswal,G., Pampell,L.R., Parish,B.J.,
        Perez,L.M., Rashid,N.D., Rives,C.M., Scherer,S.E., Shen,H.,
        Simon,M.L., Vo,Q.K., Wei,Y., Williamson,A.L., Worley,K., Zhou,X.,
        Naylor,S.L. and Gibbs,R.A.
        Direct Submission
        Unpublished
        2 (bases 1 to 123943)
        Worley,K.C.
        Direct Submission
        Submitted (11-DEC-1998) Molecular and Human Genetics, Baylor
        College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
        3 (bases 1 to 123943)
        Worley,K.C.
        Direct Submission
        Submitted (24-MAR-1999) Human Genome Sequencing Center, Department
        of Molecular and Human Genetics, Baylor College of Medicine, One
        Baylor Plaza, Houston, TX 77030, USA
        4 (bases 1 to 123943)
        Worley,K.C.
        Direct Submission
        Submitted (26-MAR-1999) Human Genome Sequencing Center, Department
        of Molecular and Human Genetics, Baylor College of Medicine, One
        Baylor Plaza, Houston, TX 77030, USA
        5 (bases 1 to 123943)
        Worley,K.C.
        Direct Submission
        Submitted (01-APR-1999) Human Genome Sequencing Center, Department
        of Molecular and Human Genetics, Baylor College of Medicine, One
        Baylor Plaza, Houston, TX 77030, USA
        On Apr 2, 1999 this sequence version replaced gi:4508049.
        INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
        gc-help@bcm.tmc.edu
CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.
ANNOTATION OF FEATURES:

```

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES	Location/Qualifiers	
	source	
repeat_region	1. .123943	/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="3p21.1-9"
		/clone="RPC14-793P23"
	2078. .2357	/rpt_family="AluSx"
		complement(4381..4674)
repeat_region		/rpt_family="AluSg"
repeat_region		complement(6668..6724)
repeat_region		/rpt_family="MIR"
repeat_region		complement(6725..7029)
repeat_region		/rpt_family="AluSc"
repeat_region		complement(7040..7197)
repeat_region		/rpt_family="MIR"
gene	Join(7498..7643,7987..8221,8767..8939)	/gene="ou12f08.x1 Homo sapiens cDNA AI004779"
STS	8793. .8921	/standard_name="G27739"
STS		/db_xref="dbSTS:41594"
STS	9133. .9283	/standard_name="D3S3975"
STS		/db_xref="dbSTS:10759"
gene	complement(Join(9446..10863,11044..11116,11221..11313,11580..11740,11858..12336,12895..13028,13552..13736,14204..14351,14692..14815,15268..15346,15613..15755,15838..15899,16397..16516,16913..17045,17993..18047,18153..18183,18281..18320))	/gene="Human mRNA for KIAA0272 gene D87462"
STS	9497. .9735	/standard_name="D3S1829"
STS		/db_xref="GDB:626512"
misc_feature	13712. .13962	/function="Low coverage"
repeat_region		complement(18321..18407)
misc_feature	18522..18742	/rpt_family="GC_rich"
repeat_region		/function="Low coverage"
repeat_region		complement(18554..18661)
gene	Join(19128..19686,21251..21360,22436..22488,22936..23027,28273..28374,28751..28875,29343..29502,30088..30166)	/gene="Unigene cluster containing AJ223811, AA280976, and AI199338"
repeat_region	20373. .20654	

Query Match 3.1%; Score 52; DB 9; Length 123943;  
Best Local Similarity 100.0%; Pred. No. 4.1e-16;

repeat_region	/rpt_family="AluJb"
repeat_region	20931. .21138
repeat_region	/rpt_family="MIR"
repeat_region	21525. .21781
repeat_region	/rpt_family="MIR"
repeat_region	21993. .22279
repeat_region	/rpt_family="AluSx"
repeat_region	23167. .23448
repeat_region	/rpt_family="AluJb"
repeat_region	complement(23639..23858)
repeat_region	/rpt_family="AluSg"
repeat_region	complement(23859..24621)
repeat_region	/rpt_family="L1ME2"
repeat_region	complement(24622..24923)
repeat_region	/rpt_family="AluY"
misc_feature	24932. .25002
misc_feature	/gene="Unigene cluster containing AJ223811, AA280976, and AI199338"
misc_feature	/function="Low coverage"
repeat_region	24993. .25290
repeat_region	/rpt_family="AluY"
repeat_region	complement(25299..25521)
repeat_region	/rpt_family="L1ME2"
repeat_region	complement(25497..26685)
repeat_region	/rpt_family="L1ME2"
repeat_region	27118. .27246
repeat_region	/rpt_family="MIR"
repeat_region	complement(27249..27544)
repeat_region	/rpt_family="AluSx"
repeat_region	27698. .27826
repeat_region	/rpt_family="MIR"
repeat_region	complement(27868..28170)
repeat_region	/rpt_family="AluY"
misc_feature	31632. .32081
misc_feature	/note="Region: qg87a10.x1 Homo sapiens cDNA AI221496"
STS	31774. .31924
STS	/standard_name="G38576"
STS	/db_xref="dbSTS:58631"
STS	31932. .32081
STS	/standard_name="G30450"
STS	/db_xref="dbSTS:44631"
STS	32599. .32898
STS	/rpt_family="AluSg"
STS	33780. .33819
STS	/rpt_family="CA)n"
STS	complement(33968..33996)
STS	/rpt_family="AT_rich"
STS	complement(34011..34082)
STS	/rpt_family="AluS"
STS	complement(34083..34121)
STS	/rpt_family="GAAA)n"
STS	complement(34123..34410)
STS	/rpt_family="AluSg"
STS	complement(34413..34697)
STS	/rpt_family="AluJb"
STS	complement(34718..35265)
STS	/rpt_family="L1MC3"
STS	complement(35266..35558)
STS	/rpt_family="AluSg"
STS	complement(35561..35738)
STS	/rpt_family="L1MD3"
STS	complement(35739..36036)
STS	/rpt_family="AluSg"
STS	complement(36038..36236)
STS	/rpt_family="L1MD3"
STS	complement(39127..39419)
STS	/rpt_family="AluSx"
STS	41664. .42236
STS	/note="Region: Unigene cluster containing W95768 and AA455370"

```

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 216 CACACCTGTATCCAGCTACTTAGGAGGCTGACGACGAGAAATTCGTTGAA 267
|||||
Db 35415 CACACCTGTATCCAGCTACTTAGGAGGCTGACGACGAGAAATTCGTTGAA 35364
|||||

RESULT 47
AC010988
LOCUS Homo sapiens BAC clone RP11-570F4 from 2, complete sequence. PRI 07-NOV-2001
DEFINITION AC010988
AC010988.9 GI:13992791
VERSION AC010988.9
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 130069)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
2 (bases 1 to 130069)
Waligorski, J., Abbott, A. and Baum, D.
The sequence of Homo sapiens BAC clone RP11-570F4
Unpublished
3 (bases 1 to 130069)
Waterston, R.H.
Direct Submission
Submitted (28-SEP-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 130069)
Waterston, R.H.
Direct Submission
Submitted (08-MAY-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 130069)
Waterston, R.
Direct Submission
Submitted (07-NOV-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On May 8, 2001 this sequence version replaced gi:13270813.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0570F04
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACE3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-526G2; the clone sequenced to the right is RP11-247H16, 200 bp overlap. Actual start of this clone is at base position 1 of RP11-570F4; actual end is at base position 45163 of RP11-247H16.

#### FEATURES

Location/Qualifiers	source
1..130069	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-570F4"
	/clone_lib="RP11-11"
221..280	/rpt_family="MIR"
331..420	/rpt_family="MIR"
843..1045	/rpt_family="MIR"
2082..2265	/rpt_family="L1"
2438..3332	/rpt_family="L1"
5055..5092	/rpt_family="MIR"
5093..5564	/rpt_family="MIR"
5565..5653	/rpt_family="MIR"
5654..5954	/rpt_family="Alu"
5931..5963	/rpt_family="AT_rich"
5955..6025	/rpt_family="MIR"
6022..6063	/rpt_family="MIR"
6149..6300	/rpt_family="AT_rich"
6149..6300	/rpt_family="MIR"
7233..7301	/rpt_family="MIR"
8085..8301	/rpt_family="MaLR"
8314..8590	/rpt_family="CR1"
8790..8882	/rpt_family="L2"
9855..9954	/rpt_family="MIR"
10031..10449	/rpt_family="MaLR"
11437..11542	/rpt_family="MIR"
11584..11604	/rpt_family="AT_rich"
11824..12135	/rpt_family="Alu"
12109..12142	/rpt_family="(A)n"
12152..12335	/rpt_family="MIR"
12376..13603	



repeat_region	/rpt_family="L1"	13363..13407
repeat_region	/rpt_family="AT_rich"	13543..13580
repeat_region	/rpt_family="A-rich"	13604..13880
repeat_region	/rpt_family="Alu"	13881..14247
repeat_region	/rpt_family="L1"	14221..15319
repeat_region	/rpt_family="L1"	15318..15829
repeat_region	/rpt_family="L1"	15835..15914
repeat_region	/rpt_family="L1"	15914..16189
repeat_region	/rpt_family="L1"	16187..17338
repeat_region	/rpt_family="L1"	17335..17826
repeat_region	/rpt_family="L1"	17450..17498
repeat_region	/rpt_family="(TA)n"	17839..18185
repeat_region	/rpt_family="L1"	18186..18359
repeat_region	/rpt_family="L1"	19455..19587
repeat_region	/rpt_family="MIR"	20570..20968
repeat_region	/rpt_family="L2"	22330..22698
repeat_region	/rpt_family="MaLR"	24292..24666
repeat_region	/rpt_family="MaLR"	24667..24999
repeat_region	/rpt_family="MaLR"	25084..25115
repeat_region	/rpt_family="(TC)n"	25678..25777
repeat_region	/rpt_family="MIR"	26008..26190
repeat_region	/rpt_family="MIR"	26881..27003
repeat_region	/rpt_family="MIR"	27358..27646
repeat_region	/rpt_family="L2"	28132..28215
repeat_region	/rpt_family="L2"	28211..28361
repeat_region	/rpt_family="MIR"	29057..29449
repeat_region	/rpt_family="MaLR"	29472..29560
repeat_region	/rpt_family="MIR"	29596..29878
repeat_region	/rpt_family="Alu"	30195..30488
repeat_region	/rpt_family="MER1_type"	30792..31102
repeat_region	/rpt_family="Alu"	31073..31106
repeat_region	/rpt_family="AT_rich"	

Query Match 3.1%; Score 52; DB 9; Length 130069;  
Best Local Similarity 100.0%; Pred.No. 4.1e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 208 TGGTGCACACACTGTAATCCAGCTACTTAGGAGCTGAGCGAGAGAAAT 259  
|||||  
DB 39182 TGGTGCACACACTGTATCCAGCTACTTAGGAGCTGAGCGAGAGAAAT 39233  
|||||

RESULT 48	AC044839
LOCUS	
DEFINITION	Homo sapiens chromosome 11, clone CTD-2210P24, complete sequence.
ACCESSION	AC044839
VERSION	AC044839.16 GI:22024599
KEYWORDS	HTG.
SOURCE	human.
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
AUTHORS	

AC044839 142323 bp DNA linear PRI 06-AUG-2002  
Homo sapiens chromosome 11, clone CTD-2210P24, complete sequence.  
AC044839.16 GI:22024599  
HTG.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Birren,B., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome 11, clone CTD-2210P24  
Unpublished  
2 (bases 1 to 142323)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,  
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Ollivar,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Willson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (12-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 142323)  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faró,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Macdonald,P., Major,J., Matthews,C.,  
Liu,G., Maclean,C., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
McCarthy,M., Melrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Willson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (05-JUL-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 142323)  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faró,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-pierre,N., Hagos,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,

Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (31-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

5 (bases 1 to 142323)

REFERENCE  
AUTHORS

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## Direct Submission

Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 31, 2002 this sequence version replaced gi:21699226.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L8906

Center clone name: 2210\_P\_24

## FEATURES

## source

Location/Qualifiers

1..142323

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="CTD-2210P24"

26...327

/rpt\_family="MER102b"

complement(590..783)

/rpt\_family="MIR"

complement(771..872)

/rpt\_family="MIR"

complement(1992..2145)

/rpt\_family="MLT1L"

complement(2514..2599)

/rpt\_family="AluSg/x"

complement(2738..2858)

/rpt\_family="AluSg/x"

complement(2864..2931)

/rpt\_family="MIR"

2964..3046

/rpt\_family="MIR"

3372..3685

/rpt\_family="AluSx"

complement(3688..4141)

repeat\_region

/rpt\_family="MER89"

4165..4322

/rpt\_family="MIR"

4347..4609

/rpt\_family="AluJo"

complement(4610..4907)

/rpt\_family="AluSg"

4945..5114

/rpt\_family="MIR3"

complement(5463..5864)

/rpt\_family="LTR30"

complement(6008..6302)

/rpt\_family="AluSx"

complement(6315..6345)

/rpt\_family="MIR"

complement(6346..6639)

/rpt\_family="AluSg"

complement(6640..6778)

/rpt\_family="MIR"

6891..7103

/rpt\_family="MIR"

7149..7293

/rpt\_family="MIR3"

complement(7349..7474)

/note="single clone coverage"

complement(7640..7644)

/note="<30 qual SINGL region"

complement(7737..7845)

/rpt\_family="MLT1H"

complement(7836..7841)

/note="<30 qual SINGL region"

complement(7912..8057)

/rpt\_family="MLT1H"

complement(8058..8364)

/rpt\_family="AluY"

complement(8365..8450)

/rpt\_family="MLT1H"

complement(8436..8442)

/note="<30 qual SINGL region"

8597..9010

/rpt\_family="L2"

9116..9173

/rpt\_family="L3"

9405..9768

/rpt\_family="MLT1A1"

9774..10289

/rpt\_family="MLT1H1"

10333..10612

/rpt\_family="L2"

10451..10513

/note="single clone coverage"

10613..10922

/rpt\_family="AluSg"

10650..10659

/note="<30 qual SINGL region"

10689..10699

Query Match

Best Local Similarity 3.1%; Score 52; DB 9; Length 142323;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 CACACCTGTATCCAGGTACTTAGGAGCTGAGCAGGAGAAATTCCTTGAA 267

|||||

Db 81052 CACACCTGTATCCAGGTACTTAGGAGCTGAGCAGGAGAAATTCCTTGAA 81103

RESULT 49

AC092338/c

LOCUS

AC092338

DEFINITION Homo sapiens chromosome 16 clone RP11-21M24, complete sequence.

AC092338

AC092338.3 GI:18481997

KEYWORDS HTG.

SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 143125)  
TITLE DOE Joint Genome Institute.  
JOURNAL Sequencing of Human Chromosome 16  
REFERENCE  
AUTHORS 2 (bases 1 to 143125)  
TITLE Direct Submission  
JOURNAL DOE Joint Genome Institute.  
REFERENCE  
AUTHORS 3 (bases 1 to 143125)  
TITLE Direct Submission  
JOURNAL DOE Joint Genome Institute.  
REFERENCE  
AUTHORS 1 (bases 1 to 143125)  
TITLE Direct Submission  
JOURNAL DOE Joint Genome Institute.  
COMMENT On Feb 3, 2002 this sequence version replaced gi:17027254.  
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.  
-----  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.  
-----  
FEATURES  
source  
1..143125  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-21M24"  
BASE COUNT 38759 a 31947 c 32716 g 39703 t  
ORIGIN  
Query Match 3.1%; Score 52; DB 9; Length 143125;  
Best Local Similarity 100.0%; Pred. No. 4e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 216 CACACCTGTAAATCCAGCTACTTAGGAGCTGAGCAGGAGGAGATTGCTTCAA 267  
|||||  
Db 50339 CACACCTGTAAATCCAGCTACTTAGGAGCTGAGCAGGAGGAGATTGCTTCAA 50288  
RESULT 50  
AC013828  
LOCUS AC013828 143577 bp DNA linear PRI 04-SEP-2001  
DEFINITION Homo sapiens chromosome , clone RP11-23B7, complete sequence.  
ACCESSION AC013828  
VERSION AC013828.10 GI:15294310  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 143577)  
TITLE Birren, B., Linton, L., Nusbaum, C., and Lander, E.  
JOURNAL Homo sapiens chromosome, clone RP11-23B7  
REFERENCE  
AUTHORS 2 (bases 1 to 143577)

## AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B., Brown, A., Castle, A., Collangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellyano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Horton, L., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Morrow, J., Navlor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

## TITLE

Direct Submission

## JOURNAL

Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## REFERENCE

3 (bases 1 to 143577)

## AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazato, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellyano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Menes, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

## TITLE

Direct Submission

## JOURNAL

Submitted (04-SEP-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 27, 2001 this sequence version replaced gi:14547834.

## COMMENT

All repeats were identified using RepeatMasker:

Smit, A.F.A. &amp; Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L4171

Center clone name: 23\_B\_7

-----

## FEATURES

Location/Qualifiers

Source

1..143577

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="RP11-23B7"

/clone\_lib="RP11-23B7"

complement(670..775)

/rpt\_family="MIR3"

complement(2047..2329)

/rpt\_family="AluY"

2876..2896

/rpt\_family="AT-rich"

complement(4044..4169)

/rpt\_family="FLAM\_A"

complement(4199..4627)

/rpt\_family="MSTB1"

complement(4831..4872)

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

```
repeat_region /rpt_family="MER61B"
5133. .5453
repeat_region /rpt_family="MLTIG"
5495. .5623
repeat_region /rpt_family="MLTIG"
5669. .5690
repeat_region /rpt_family="AT_rich"
6138. .6346
repeat_region /rpt_family="MER58A"
complement(6378. .6689)
repeat_region /rpt_family="AluJb"
6889. .7187
repeat_region /rpt_family="AluSx"
complement(7321. .7456)
repeat_region /rpt_family="FLAM_C"
complement(8235. .8396)
repeat_region /rpt_family="MIR"
complement(9700. .9762)
repeat_region /rpt_family="MIR3"
complement(10332. .10435)
repeat_region /rpt_family="HY1"
complement(10737. .10822)
repeat_region /rpt_family="MIR"
11523. .11887
repeat_region /rpt_family="L1PA4"
12381. .12426
repeat_region /rpt_family="L2"
complement(12600. .12711)
repeat_region /rpt_family="FLAM_A"
12807. .12919
repeat_region /rpt_family="MER94"
complement(12954. .13651)
repeat_region /rpt_family="L1ME3"
complement(13652. .13954)
repeat_region /rpt_family="AluSx"
complement(13955. .14704)
repeat_region /rpt_family="L1ME3"
14729. .14749
repeat_region /rpt_family="AT_rich"
complement(14786. .15179)
repeat_region /rpt_family="MLT1A1"
complement(15180. .15601)
repeat_region /rpt_family="L1MB5"
15602. .15665
repeat_region /rpt_family="MER46A"
15666. .15954
repeat_region /rpt_family="AluSx"
15955. .16120
repeat_region /rpt_family="MER46A"
complement(16121. .16340)
repeat_region /rpt_family="L1MB5"
complement(16665. .16744)
repeat_region /rpt_family="HAL1"
16874. .17072
repeat_region /rpt_family="MER30"
complement(18234. .18318)
repeat_region /rpt_family="MIR3"
complement(19051. .19186)
repeat_region /rpt_family="MIR"
19699. .19730
repeat_region /rpt_family="AT_rich"
complement(20696. .20883)
repeat_region /rpt_family="MER20"
20956. .21257
repeat_region /rpt_family="AluSx"
21773. .21963
repeat_region /rpt_family="L2"
complement(22007. .22121)
repeat_region /rpt_family="L2"
22261. .22329
repeat_region /rpt_family="L2"
22330. .22623
repeat_region /rpt_family="AluY"
```

```
repeat_region 22624. .22847
/rpt_family="L2"
repeat_region 23118. .23156
/rpt_family="AT_rich"
repeat_region 23800. .23931
/rpt_family="MIR"
repeat_region 24755. .24780
/rpt_family="AT_rich"
repeat_region 25038. .25201
/rpt_family="L1MASA"
repeat_region 25239. .25311
/rpt_family="AluJb"
repeat_region 25312. .25344
/rpt_family="AT_rich"
repeat_region 25627. .25769
/rpt_family="MIR"
repeat_region 26152. .26172
/rpt_family="AT_rich"
repeat_region complement(26951. .27162)
/rpt_family="AluJb"
repeat_region 27540. .27842
/rpt_family="AluSx"
repeat_region complement(28209. .28402)
/rpt_family="MER5A"
repeat_region complement(28406. .28465)
/rpt_family="MER5A"
repeat_region complement(28591. .28742)
/rpt_family="MER5A"
repeat_region complement(30280. .30490)
/rpt_family="L1MA8"
repeat_region complement(30558. .30781)
/rpt_family="MIR"
repeat_region complement(30861. .31046)
/rpt_family="L1MC/D"
repeat_region 31425. .31461
/rpt_family="AT_rich"
repeat_region 31902. .32160
/rpt_family="AluSx"
repeat_region 32226. .32259
/rpt_family="(CA)n"
```

Query Match 3.1%; Score 52; DB 9; Length 143577;  
Best Local Similarity 100.0%; Pred. No. 4e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 CACACCTGTAATCCAGCTACTTAGGAGCTCAGGAGAGAAATGCTTGAA 267  
|||||  
Db 58873 CACACCTGTAATCCAGCTACTTAGGAGCTCAGGAGAGAAATGCTTGAA 58924  
|||||

Search completed: February 11, 2003, 13:33:34  
Job time : 8121 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 05:35:06 : Search time 373 Seconds  
(without alignments)  
10143.049 Million cell updates/sec

Title: US-09-942-310-2  
Perfect score: 1680  
Sequence: 1 gaattcaagaccagctgga.....catcttcctgctcgtgtgg 1680

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 20  
Total number of hits satisfying chosen parameters: 93769

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N_Geneseq_101002:*			
1:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*		
2:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*		
3:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*		
4:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*		
5:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*		
6:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*		
7:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*		
8:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*		
9:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*		
10:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*		
11:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*		
12:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*		
13:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*		
14:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*		
15:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*		
16:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*		
17:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*		
18:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*		
19:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*		
20:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*		
21:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*		
22:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*		
23:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*		
24:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1680	100.0	1680	24	AAD34214 Human CYP2D6 gene
2	1536	90.2	1669	22	AH26169 Human cytochrome P
3	1480	88.1	1669	22	AH26179 Human cytochrome P
4	1338	79.6	9432	24	AAD34213 Human cytochrome P
5	907	54.0	6472	24	ABQ72364 Human CYP2D6 gene,
6	660	39.3	6472	24	ABQ72315 Human CYP2D6 gene,
7	138	8.2	1593	22	AD00937 Human drug metabol
8	61	3.6	1494	16	AQ87729 Human auxillary cy
9	61	3.6	1494	16	AQ87730 Human auxillary cy



C 229	38	2.3	18564	22	AAK65421	Human immune/haema	C 302	37	2.2	16892	22	AAK90086	Human digestive sy
C 230	38	2.3	18564	22	AAK84438	Human immune/haema	C 303	37	2.2	17070	22	AAK80632	Human immune/haema
C 231	38	2.3	21742	21	AAF20938	Human high affinity	C 304	37	2.2	18488	22	AAK83173	Human immune/haema
C 232	38	2.3	21742	21	AAK34816	Human adenosine re	C 305	37	2.2	21332	22	AAK74878	Human immune/haema
C 233	38	2.3	26048	22	AAK36056	Human cardiovascular	C 306	37	2.2	21332	22	AAK83177	Human immune/haema
C 234	38	2.3	33030	22	AAF29337	Atopy related gene	C 307	37	2.2	22481	17	AAT11658	PEDF full length s
C 235	38	2.3	44100	21	ABN97975	Human retroviral s	C 308	37	2.2	22484	24	ABO88196	Human osteoblast d
C 236	38	2.3	117609	21	AAF21435	Human receptor-rel	C 309	37	2.2	22484	24	ABN96844	Gene #3342 used to
C 237	38	2.3	143068	21	AAF211702	Human low adenosin	C 310	37	2.2	22484	24	ABA98882	Nucleotide sequenc
C 238	38	2.3	143068	21	AAF211702	Human low adenosin	C 311	37	2.2	22484	24	ABA05882	Angiogenesis inhib
C 239	38	2.3	143068	21	AAK34983	Human adenosine re	C 312	37	2.2	24102	22	AAK74877	Human immune/haema
C 240	38	2.3	143068	21	AAK35150	Human adenosine re	C 313	37	2.2	24964	24	ABK6595	Human SA (rat hype
C 241	38	2.3	143068	24	ABU68124	Ovary cancer relat	C 314	37	2.2	24999	22	AAK65622	Human immune/haema
C 242	38	2.3	149412	21	AAK35151	Human adenosine re	C 315	37	2.2	26496	22	ABA17696	Human nervous syst
C 243	38	2.3	152740	21	AAF21273	Human low adenosin	C 316	37	2.2	26657	24	AAI72317	Human transporter
C 244	38	2.3	167343	24	ABL64403	Stomach cancer rel	C 317	37	2.2	30360	22	AAK82759	Human immune/haema
C 245	38	2.3	167343	24	ABL67239	Thyroid cancer rel	C 318	37	2.2	31405	22	AAK67293	Human immune/haema
C 246	38	2.3	169739	24	AQ888186	Human osteoblast d	C 319	37	2.2	31405	22	AAK74865	Human immune/haema
C 247	37	2.2	178	21	AAC13748	Human secreted pro	C 320	37	2.2	31405	22	AAK83153	Human immune/haema
C 248	37	2.2	293	22	AAL35034	Human musculoskele	C 321	37	2.2	32189	22	AAL04670	Human reproductive
C 249	37	2.2	361	21	AAC14976	Human secreted pro	C 322	37	2.2	32189	23	ABL97577	Human testicular a
C 250	37	2.2	362	23	ABV15223	Human prostate exp	C 323	37	2.2	32221	22	AAK539766	Genomic sequence #
C 251	37	2.2	380	21	AAH30221	Human colon cancer	C 324	37	2.2	32221	22	AAK90119	Human digestive sy
C 252	37	2.2	385	23	ABV06051	Human prostate exp	C 325	37	2.2	33147	22	AAK67282	Human immune/haema
C 253	37	2.2	390	23	ABV19489	Human prostate exp	C 326	37	2.2	35465	22	AAF54723	Nucleotide sequenc
C 254	37	2.2	410	21	AAC13245	Human secreted pro	C 327	37	2.2	35959	22	AAK78275	Human immune/haema
C 255	37	2.2	429	22	AAL03690	Human reproductive	C 328	37	2.2	41684	21	AAA28150	Human purh gene ge
C 256	37	2.2	438	24	ABL82984	Human ovarian canc	C 329	37	2.2	51469	22	AAK69322	Human immune/haema
C 257	37	2.2	446	22	AAK78049	Human immune/haema	C 330	37	2.2	51469	22	AAK70270	Human immune/haema
C 258	37	2.2	452	23	ABV36022	Human prostate exp	C 331	37	2.2	51469	22	AAK78813	Human immune/haema
C 259	37	2.2	452	23	ABV45082	Human prostate exp	C 332	37	2.2	68356	22	AAK67283	Human immune/haema
C 260	37	2.2	552	22	AAK78048	Human immune/haema	C 333	37	2.2	68356	22	AAK83212	Human immune/haema
C 261	37	2.2	621	22	AAK74881	Human immune/haema	C 334	37	2.2	80240	20	AAV83940	Human immune/haema
C 262	37	2.2	630	24	ABN61162	Human cancer relat	C 335	37	2.2	80595	20	AAV83939	NC-contig derived
C 263	37	2.2	643	23	ABK42137	Genomic sequence #	C 336	37	2.2	110608	24	ABK83572	Human CDNA differe
C 264	37	2.2	954	22	AAK74882	Human immune/haema	C 337	37	2.2	162450	21	AAZ86967	Retinoblastoma bin
C 265	37	2.2	1174	22	AAL36246	Human musculoskele	C 338	37	2.2	174424	24	ABL68122	Ovary cancer relat
C 266	37	2.2	2421	22	AAI69607	Human RNA helicase	C 339	37	2.2	183999	22	AAF92831	Human ABC1 genomic
C 267	37	2.2	3200	24	ABK47724	Genomic sequence e	C 340	36	2.1	132	22	AAK534599	Human DNA for a no
C 268	37	2.2	4354	22	AAK83825	Human immune/haema	C 341	36	2.1	132	22	AAK534600	Human DNA for a no
C 269	37	2.2	4375	24	ABK24527	EIF-2alpha kinase	C 342	36	2.1	320	22	ABA17067	Human nervous syst
C 270	37	2.2	5668	22	AAK32797	Human genomic DNA	C 343	36	2.1	323	24	ABL86266	Human ovarian canc
C 271	37	2.2	5771	22	AAK75155	Human immune/haema	C 344	36	2.1	328	21	AAK05114	Human secreted pro
C 272	37	2.2	5797	22	ABA15723	Human nervous syst	C 345	36	2.1	358	22	AAK538359	Novel human diagno
C 273	37	2.2	5797	22	AAL03405	Human reproductive	C 346	36	2.1	363	22	AAI80917	Human polynucleoti
C 274	37	2.2	6496	22	AAK72255	Human immune/haema	C 347	36	2.1	374	23	ABV11033	Human prostate exp
C 275	37	2.2	6751	22	AAK83163	Human immune/haema	C 348	36	2.1	378	23	ABV01864	Human prostate exp
C 276	37	2.2	6767	22	AAK83147	Human immune/haema	C 349	36	2.1	382	22	AAH33056	Human colon cancer
C 277	37	2.2	6943	22	AAK83198	Human immune/haema	C 350	36	2.1	384	24	ABL84741	Human ovarian canc
C 278	37	2.2	7906	22	AAK83154	Human immune/haema	C 351	36	2.1	391	22	AAI99228	Human excretory re
C 279	37	2.2	8040	22	AAL05543	Human immune/haema	C 352	36	2.1	391	22	AAI63578	Human kidney relat
C 280	37	2.2	9133	22	AAK83165	Human reproductive	C 353	36	2.1	393	22	AAF66809	Novel human polynu
C 281	37	2.2	9203	22	AAK87134	Human immune/haema	C 354	36	2.1	415	23	AAV32181	Human prostate exp
C 282	37	2.2	9358	24	ABK84319	Human CDNA differe	C 355	36	2.1	415	23	ABV41112	Human prostate exp
C 283	37	2.2	9358	24	ABL61905	Colon adenocarcino	C 356	36	2.1	456	23	ABV52764	Human prostate exp
C 284	37	2.2	9382	22	AAK83150	Human immune/haema	C 357	36	2.1	471	24	ABL77841	Human ovarian canc
C 285	37	2.2	10497	22	AAK67302	Human immune/haema	C 358	36	2.1	527	22	AAI88189	Human polynucleoti
C 286	37	2.2	10497	22	AAK83166	Human immune/haema	C 359	36	2.1	543	23	ABV58687	Human prostate exp
C 287	37	2.2	11428	22	AAK83156	Human immune/haema	C 360	36	2.1	545	22	AAH10383	Human CDNA clone (
C 288	37	2.2	11469	22	AAK36426	Human musculoskele	C 361	36	2.1	554	22	AAH13313	Human CDNA clone (
C 289	37	2.2	11744	22	AAK83160	Human immune/haema	C 362	36	2.1	589	23	ABV58970	Human prostate exp
C 290	37	2.2	11853	22	AAK67297	Human immune/haema	C 363	36	2.1	747	22	AAL24339	Human breast canc
C 291	37	2.2	11853	22	AAK83162	Human immune/haema	C 364	36	2.1	773	24	ABO89812	Human prostate exp
C 292	37	2.2	13774	22	AAK83209	Human immune/haema	C 365	36	2.1	807	24	ABQ89867	Human prostate exp
C 293	37	2.2	14258	22	AAK83210	Human immune/haema	C 366	36	2.1	811	24	ABQ89853	Human prostate exp
C 294	37	2.2	14260	22	AAK83183	Human immune/haema	C 367	36	2.1	1017	22	AAK72092	Human immune/haema
C 295	37	2.2	14260	22	AAK83217	Human immune/haema	C 368	36	2.1	1017	22	AAK72093	Human immune/haema
C 296	37	2.2	15201	22	AAK83176	Human immune/haema	C 369	36	2.1	1282	24	AAK562532	CDNA sequence #319
C 297	37	2.2	15364	22	AAK83152	Human immune/haema	C 370	36	2.1	1862	22	AAK32779	Human secreted pro
C 298	37	2.2	15772	22	AAK83220	Human immune/haema	C 371	36	2.1	1868	22	AAH17957	Human CDNA sequenc
C 299	37	2.2	16892	22	AAK39745	Genomic sequence #	C 372	36	2.1	1891	24	AAL46222	Human liver cancer
C 300	37	2.2	16892	22	AAK80873	Human immune/haema	C 373	36	2.1	2609	22	AAH17513	Human CDNA sequenc
C 301	37	2.2	16892	22	AAK89201	Human digestive sy	C 374	36	2.1	3159	22	AAK80811	Human immune/haema

375	36	2.1	3842	24	ABK54438	Human BMPR2 partia	448	35	2.1	333	22	AAS38577	Novel human diagno
376	36	2.1	4457	22	AAL37317	Human musculoskele	449	35	2.1	339	24	ABL86365	Human ovarian canc
377	36	2.1	5069	22	AAK67404	Human immune/haema	450	35	2.1	339	22	AAH70198	Human cervical can
378	36	2.1	5243	22	AAH18240	Human CDNA sequenc	451	35	2.1	371	22	AAI90990	Human polynucleoti
379	36	2.1	5576	22	AAK72826	Human immune/haema	452	35	2.1	374	22	AAS38576	Novel human diagno
380	36	2.1	5576	22	AAK72901	Human immune/haema	453	35	2.1	376	22	AAF65548	Novel human polynu
381	36	2.1	5580	22	AAK72824	Human immune/haema	454	35	2.1	377	21	AAC26130	Human secreted pro
382	36	2.1	5580	22	AAK72899	Human immune/haema	455	35	2.1	378	24	ABL84443	Human ovarian canc
383	36	2.1	6170	22	AAK74779	Human immune/haema	456	35	2.1	386	22	AAI81243	Human polynucleoti
384	36	2.1	6202	22	AAK85135	Human immune/haema	457	35	2.1	386	22	AAK57679	Human immune/haema
385	36	2.1	6868	22	AAK87324	Human immune/haema	458	35	2.1	386	24	ABL70085	Pancreas cancer re
386	36	2.1	6896	22	AAK69978	Human immune/haema	459	35	2.1	388	24	ABL85199	Human ovarian canc
387	36	2.1	9840	22	AAK70512	Human immune/haema	460	35	2.1	394	22	AAI92798	Human polynucleoti
388	36	2.1	10198	22	AAI06055	Human reproductive	461	35	2.1	397	24	ABK46028	CDNA encoding colo
389	36	2.1	10198	22	AAI06056	Human reproductive	462	35	2.1	411	22	AAI87862	Human polynucleoti
390	36	2.1	10198	23	ABL98620	Human testicular a	463	35	2.1	412	22	AAK70876	Human polynucleoti
391	36	2.1	10198	23	ABL98621	Human testicular a	464	35	2.1	412	22	AAK70876	Human immune/haema
392	36	2.1	11726	22	ABAL5436	Human nervous syst	465	35	2.1	450	22	AAF65537	Novel human polynu
393	36	2.1	11901	20	AAK02998	Human IL-1ra BAC c	466	35	2.1	451	24	ABL87490	Human cDNA encodin
394	36	2.1	12730	22	AAK70514	Human immune/haema	467	35	2.1	468	22	AAI88978	Human ovarian canc
395	36	2.1	13165	22	ABAL6856	Human pancreatic c	468	35	2.1	471	24	ABL83804	Human polynucleoti
396	36	2.1	13165	22	AAK89348	Human digestive sy	469	35	2.1	494	22	AAI14931	Human breast canc
397	36	2.1	13345	22	AAK66901	Human immune/haema	470	35	2.1	515	22	AAI15384	Human breast canc
398	36	2.1	17687	22	AAS42069	Genomic sequence #	471	35	2.1	515	22	AAI24230	Human breast canc
399	36	2.1	19616	22	ABAL6094	Human nervous syst	472	35	2.1	536	23	AAS78843	DNA encoding novel
400	36	2.1	19616	22	ABAL6856	Human nervous syst	473	35	2.1	538	23	ABV51846	Human prostate exp
401	36	2.1	19616	22	AAI36855	Human nervous syst	474	35	2.1	547	24	ABQ57021	Human colon cancer
402	36	2.1	19815	22	AAS42064	Genomic sequence #	475	35	2.1	548	22	AAH16890	Human cDNA clone (
403	36	2.1	20522	22	AAI37409	Human musculoskele	476	35	2.1	562	22	AAI14786	Human breast canc
404	36	2.1	22635	22	AAI07203	Human musculoskele	477	35	2.1	576	22	AAH09266	Human cDNA clone (
405	36	2.1	23071	22	AAS26699	Human reproductive	478	35	2.1	599	24	ABN60256	Human prostate exp
406	36	2.1	28818	22	AAI35901	Human genomic DNA	479	35	2.1	600	22	AAI23789	Human cancer relat
407	36	2.1	28974	22	AAS32821	Human musculoskele	480	35	2.1	604	24	ABN60371	Human breast relat
408	36	2.1	28974	22	AAK69977	Human immune/haema	481	35	2.1	608	23	ABV54755	Human prostate exp
409	36	2.1	30013	22	AAI36932	Human immune/haema	482	35	2.1	613	22	ABA20704	Human nervous syst
410	36	2.1	30013	22	AAS41360	Genomic sequence #	483	35	2.1	613	22	ABV58091	Human nervous syst
411	36	2.1	32216	22	ABAO7741	Human ovarian and	484	35	2.1	614	22	AAK20703	Human prostate exp
412	36	2.1	32216	22	AAI06925	Human reproductive	485	35	2.1	620	22	AAK77824	Human nervous syst
413	36	2.1	34488	22	AAI06925	Human neuroblastom	486	35	2.1	620	22	AAK77825	Human immune/haema
414	36	2.1	48037	22	AAK84729	Human immune/haema	487	35	2.1	641	22	AAK66085	Human immune/haema
415	36	2.1	48045	22	AAK85983	Human immune/haema	488	35	2.1	641	22	AAK66085	Human immune/haema
416	36	2.1	48045	22	AAK84730	Human immune/haema	489	35	2.1	662	22	AAK69380	Human immune/haema
417	36	2.1	48045	22	AAK85984	Human immune/haema	490	35	2.1	668	22	AAI13388	Human immune/haema
418	36	2.1	49744	24	ABK12807	Human tumour suppr	491	35	2.1	686	22	AAI22343	Human breast canc
419	36	2.1	52354	24	ABQ35032	Human transporter	492	35	2.1	700	22	AAH93002	Human breast relat
420	36	2.1	92638	24	ABQ88096	Human osteoblast d	493	35	2.1	710	23	AAS78555	DNA encoding novel
421	36	2.1	99014	24	ABN96931	Gene #3429 used to	494	35	2.1	716	22	AAI22258	Human inflammatory
422	36	2.1	109906	24	ABK94411	DNA encoding endot	495	35	2.1	725	22	AAK67206	Human breast canc
423	36	2.1	112460	24	ABK83567	Human CDNA differe	496	35	2.1	744	20	AAZ15646	Human immune/haema
424	36	2.1	121162	21	AAK66548	Human kinesin-like	497	35	2.1	746	22	AAI13473	Human gene express
425	36	2.1	138169	21	AAA34791	Human adenosine re	498	35	2.1	773	22	AAI95087	Human breast relat
426	36	2.1	141589	21	AAI20913	Human ELAM-1 polyn	499	35	2.1	803	22	AAI23648	Human neuroblastom
427	36	2.1	141589	21	AAI21127	Human low adenosin	500	35	2.1	819	22	AAH05049	Human breast relat
428	36	2.1	141589	21	AAI21152	Human low adenosin	501	35	2.1	837	24	ABQ89125	Human CDNA clone (
429	36	2.1	141589	21	AAA35005	Human adenosine re	502	35	2.1	837	22	AAI36591	Human prostate exp
430	36	2.1	141589	21	AAA35030	Human adenosine re	503	35	2.1	903	22	AAI36591	Human muscleoskele
431	36	2.1	146981	21	AAI21442	Human ELAM-1 polyn	504	35	2.1	981	24	ABK95691	Human cytokine rec
432	36	2.1	154465	24	AAD28763	Human AKAP allelic	505	35	2.1	999	22	AAI84751	Nucleotide sequenc
433	36	2.1	158245	24	AAD28762	Human AKAP allelic	506	35	2.1	1012	22	AAK83000	Human immune/haema
434	36	2.1	161425	22	AAH02340	Human AKAP allelic	507	35	2.1	1274	22	ABA20573	Human nervous syst
435	36	2.1	162025	22	AAH02339	Human AKAP10 gene	508	35	2.1	1274	22	ABA20573	Human nervous syst
436	36	2.1	162025	22	AAD28758	Human AKAP allelic	509	35	2.1	1438	20	AAI23317	Human l-alpha-Ohas
437	36	2.1	162025	24	AAD28759	Human hypcretin r	510	35	2.1	1494	22	AAH17093	Human cDNA sequenc
438	36	2.1	168575	22	AAH21613	Human ABC1 genomic	511	35	2.1	1608	20	AAZ07544	Human lalpha hydro
439	36	2.1	183999	22	AAI92831	Human factor-relat	512	35	2.1	1674	22	AAI89482	Human polynucleoti
440	36	2.1	209273	21	AAI21437	Human CDNA differe	513	35	2.1	1779	22	ABA19142	Human nervous syst
441	36	2.1	220895	24	ABK84798	Human SNP oligonuc	514	35	2.1	1779	22	AAS28340	Genomic sequence #
442	35	2.1	51	22	AAI31460	Human immune/haema	515	35	2.1	1780	22	ABA19141	Human nervous syst
443	35	2.1	129	20	AAK65422	Human gene express	516	35	2.1	1780	22	AAS28339	Genomic sequence #
444	35	2.1	309	24	ABL86258	Human ovarian canc	517	35	2.1	1842	22	AAI15687	Human cDNA sequenc
445	35	2.1	312	24	ABL87771	Human ovarian canc	518	35	2.1	1943	22	AAI60162	Human polynucleoti
446	35	2.1	312	24	ABL87771	Human ovarian canc	519	35	2.1				
447	35	2.1	321	24	ABL87757	Human ovarian canc	520	35	2.1				



521	35	2.1	1986	22	AAF26548	DNA encoding human	594	35	2.1	30967	17	AAT32454	Calpain large subu
522	35	2.1	1988	22	AAI199239	Human excretory re	595	35	2.1	31885	22	ABA19143	Human nervous syst
523	35	2.1	1988	22	AAS36288	Human cardiovascular	c 596	35	2.1	31885	22	AAL36165	Human musculoskele
524	35	2.1	1988	22	AAI63589	Human kidney relat	c 597	35	2.1	31885	22	AAL36176	Human musculoskele
525	35	2.1	2189	22	AAH17096	Human cDNA sequenc	c 598	35	2.1	31885	22	AAL04523	Human reproductive
526	35	2.1	2211	22	AAK71890	Human immune/haema	c 599	35	2.1	31885	22	AAS28341	Genomic sequence #
527	35	2.1	2440	22	AAH17713	Human cDNA sequenc	c 600	35	2.1	31885	23	ABL97447	Human testicular a
528	35	2.1	2515	22	AAH17444	Human granulocyte	c 601	35	2.1	32169	22	ABA14358	Human nervous syst
529	35	2.1	2660	24	AAD28060	Human adenyllyl and	c 602	35	2.1	32190	22	AAK89689	Human digestive sy
530	35	2.1	3073	24	ABK69999	cDNA encoding huma	c 603	35	2.1	32207	22	AAL36885	Human musculoskele
531	35	2.1	4233	22	AAK65688	Human immune/haema	c 604	35	2.1	32207	22	AAI04278	Human reproductive
532	35	2.1	4352	22	ABA19615	Human nervous syst	605	35	2.1	39265	24	ABL52838	Poly nucleotide seq
533	35	2.1	4743	22	AAL05210	Human reproductive	606	35	2.1	41159	22	AAK56331	Human immune/haema
534	35	2.1	4743	23	ABL98093	Human testicular a	c 607	35	2.1	43069	21	AAZ36335	Genomic sequence o
535	35	2.1	4747	22	AAI05209	Human reproductive	608	35	2.1	43950	24	AAD36022	Human kinase genom
536	35	2.1	4747	23	ABL98092	Human testicular a	c 609	35	2.1	44147	24	ABK84481	Human cDNA differe
537	35	2.1	4958	22	AAK67565	Human immune/haema	610	35	2.1	44354	22	AAK77833	Human immune/haema
538	35	2.1	4958	22	AAK68610	Human immune/haema	611	35	2.1	44354	22	AAK77836	Human immune/haema
539	35	2.1	5281	22	AAI04261	Human reproductive	612	35	2.1	44354	22	AAK77837	Human immune/haema
540	35	2.1	5491	22	AAK69044	Human immune/haema	c 613	35	2.1	46107	22	AAK71730	Human immune/haema
541	35	2.1	6040	22	AAS28595	Genomic sequence #	614	35	2.1	48436	24	ABN89533	Human corneal N-ac
542	35	2.1	6235	21	AAZ29169	Human G-CSF genom	615	35	2.1	50000	24	ABA98944	Human asthma-associ
543	35	2.1	6427	22	AAK91387	Human digestive sy	c 616	35	2.1	52845	22	AAK71437	Human immune/haema
544	35	2.1	6427	22	AAI57755	Human colorectal c	617	35	2.1	54108	24	ABK22782	Human high bone ma
545	35	2.1	6470	22	AAI58376	Human polynucleoti	618	35	2.1	57273	24	ABK22784	Human high bone ma
546	35	2.1	6544	23	ABK42657	Genomic sequence #	c 619	35	2.1	65608	24	ABL62910	Stomach cancer rela
547	35	2.1	6554	22	ABA18544	Human nervous syst	c 620	35	2.1	65608	24	ABL64414	Breast cancer relat
548	35	2.1	6554	22	ABA18924	Human nervous syst	c 621	35	2.1	65608	24	ABL67668	Oesophagus cancer
549	35	2.1	6567	22	ABA18923	Human nervous syst	622	35	2.1	66933	22	ABA82625	Human HBM gene reg
550	35	2.1	6679	21	AAZ23170	Partial sequence o	623	35	2.1	72049	22	ABA82623	Human HBM gene reg
551	35	2.1	6679	22	AAI17443	Human granulocyte	c 624	35	2.1	81001	22	AAF30035	Human apolipoprote
552	35	2.1	6892	22	AAI37405	Human musculoskele	c 625	35	2.1	90220	24	ABK83576	Human cDNA differe
553	35	2.1	6955	22	AAK79747	Human immune/haema	c 626	35	2.1	109906	24	ABK94411	DNA encoding endot
554	35	2.1	7299	22	ABA16794	Human nervous syst	627	35	2.1	114793	22	AAD08215	Human genome from
555	35	2.1	7299	22	AAK77941	Human immune/haema	c 628	35	2.1	114793	22	AAD08215	Human genome from
556	35	2.1	7426	24	ABK46568	DNA encoding G-pro	c 629	35	2.1	122888	24	ABK83569	Human cDNA differe
557	35	2.1	9440	22	AAI02976	Human reproductive	c 630	35	2.1	149671	24	ABK84797	Human cDNA differe
558	35	2.1	9878	22	ABA16974	Human nervous syst	631	35	2.1	160552	22	AAD02697	Human glycosyl sul
559	35	2.1	9883	22	ABA16973	Human nervous syst	632	35	2.1	169739	24	ABQ88186	Human osteoblast d
560	35	2.1	10210	24	AAS18100	Human angiotensin	c 633	35	2.1	172570	24	ABQ88207	Human osteoblast d
561	35	2.1	10663	21	AAI12624	Genomic DNA encodi	c 634	35	2.1	175737	24	ABK83571	Human cDNA differe
562	35	2.1	11234	22	ABA20857	Human nervous syst	635	35	2.1	325791	22	AAS43104	Human Oestrogen re
563	35	2.1	13274	22	AAL36063	Human musculoskele	636	35	2.0	51	22	AAI79827	Human nonconservat
564	35	2.1	13953	24	AAD34465	Human phospholipas	c 637	34	2.0	120	21	AAI11854	Human secreted pro
565	35	2.1	14925	22	AAK67389	Human immune/haema	c 638	34	2.0	193	22	ABA21487	Human nervous syst
566	35	2.1	16106	22	AAI35957	Human musculoskele	639	34	2.0	214	24	ABL86632	Human ovarian canc
567	35	2.1	16106	22	AAK67366	Human immune/haema	c 640	34	2.0	224	22	ABA21488	Human nervous syst
568	35	2.1	17000	24	AAI40299	Caspase 6 antisens	c 641	34	2.0	224	22	ABA21491	Human nervous syst
569	35	2.1	17794	22	ABA20310	Human nervous syst	c 642	34	2.0	224	22	ABA21493	Human nervous syst
570	35	2.1	18595	22	AAS33411	DNA encoding human	c 643	34	2.0	260	24	ABK39170	cDNA encoding lung
571	35	2.1	19846	22	AAS36182	Human cardiovascular	644	34	2.0	293	22	AAS35613	Human cardiovascular
572	35	2.1	19846	22	AAS36183	Human cardiovascular	c 645	34	2.0	305	22	ABA15810	Human nervous syst
573	35	2.1	20067	22	AAK66735	Human immune/haema	c 646	34	2.0	305	22	ABA18642	Human nervous syst
574	35	2.1	20068	22	AAK66733	Human immune/haema	647	34	2.0	325	21	AAC32103	Human secreted pro
575	35	2.1	20323	22	AAK66731	Human immune/haema	c 648	34	2.0	354	23	ABV53871	Human prostate exp
576	35	2.1	20769	22	ABA15142	Human immune/haema	c 649	34	2.0	356	24	ABL80348	Human ovarian canc
577	35	2.1	20778	22	AAK79819	Human nervous syst	650	34	2.0	361	22	AAI87213	Human polynucleoti
578	35	2.1	20911	22	AAK76454	Human immune/haema	c 651	34	2.0	378	24	ABL77884	Human ovarian canc
579	35	2.1	23934	22	ABA19145	Human nervous syst	652	34	2.0	386	22	AAI93288	Human polynucleoti
580	35	2.1	23934	22	AAL36171	Human musculoskele	c 653	34	2.0	388	22	AAI83815	Human polynucleoti
581	35	2.1	23934	22	AAI36179	Human musculoskele	654	34	2.0	404	22	AAI87929	Human polynucleoti
582	35	2.1	23934	22	AAL04522	Human reproductive	c 655	34	2.0	407	22	AAF67342	Novel human polynu
583	35	2.1	23934	22	AAS28343	Genomic sequence #	656	34	2.0	412	23	ABV19555	Human prostate exp
584	35	2.1	23934	22	AAK71442	Human immune/haema	657	34	2.0	414	22	AAI83821	Human polynucleoti
585	35	2.1	23934	23	ABL97446	Human testicular a	658	34	2.0	417	21	AAI10003	Human secreted pro
586	35	2.1	28895	22	AAK77832	Human immune/haema	c 659	34	2.0	423	22	AAI92131	Human polynucleoti
587	35	2.1	27118	22	ABA08223	Human ovarian and	660	34	2.0	433	23	ABV16862	Human prostate exp
588	35	2.1	27118	22	AAI07542	Human reproductive	661	34	2.0	441	22	AAK64882	Human immune/haema
589	35	2.1	28091	22	ABA15375	Human nervous syst	c 662	34	2.0	444	22	AAI92463	Human polynucleoti
590	35	2.1	28091	22	AAK77432	Human immune/haema	c 663	34	2.0	457	23	ABV09165	Human polynucleoti
591	35	2.1	28120	22	ABA15379	Human nervous syst	c 664	34	2.0	468	23	ABV46659	Human prostate exp
592	35	2.1	28120	22	AAK77434	Human immune/haema	665	34	2.0	470	22	AAI93260	Human polynucleoti
593	35	2.1	29629	24	ABL58699	Human kinase encod	666	34	2.0	475	22	AAH68754	Human cervical can

c 667	34	2.0	482	21	AAC24344	Human secreted pro	c 740	34	2.0	1529	22	AAS28463	Genomic sequence #
c 668	34	2.0	488	23	ABV49323	Human prostate exp	c 741	34	2.0	1529	22	AAS28464	Genomic sequence #
c 669	34	2.0	492	24	ABN64614	Human cancer relat	c 742	34	2.0	1529	22	AAS28465	Genomic sequence #
c 670	34	2.0	493	22	ABAI16673	Human nervous syst	c 743	34	2.0	1529	22	AAS28466	Genomic sequence #
c 671	34	2.0	499	22	ABAI13243	Human cDNA clone (	c 744	34	2.0	1534	20	AAX97923	Human secreted pro
c 672	34	2.0	513	22	ABAI2008	Human nervous syst	c 745	34	2.0	1544	21	AAX74341	Human secreted pro
c 673	34	2.0	514	23	ABV55804	Human prostate exp	c 746	34	2.0	1656	22	AAI80590	Human polynucleoti
c 674	34	2.0	515	22	AAH10511	Human cDNA clone (	c 747	34	2.0	1656	22	AAH18017	Human cDNA sequenc
c 675	34	2.0	516	23	ABV25904	Human prostate exp	c 748	34	2.0	1661	22	AAK68994	Human immune/haema
c 676	34	2.0	516	23	ABV25905	Human prostate exp	c 749	34	2.0	1721	22	AAH13803	Human cDNA sequenc
c 677	34	2.0	517	23	ABV29948	Human prostate exp	c 750	34	2.0	2031	22	AAF93847	Human cDNA encodin
c 678	34	2.0	520	23	ABV24067	Human prostate exp	c 751	34	2.0	2106	22	AAH17107	Human cDNA sequenc
c 679	34	2.0	524	22	AAK73289	Human immune/haema	c 752	34	2.0	2221	22	AAH16381	Human cDNA sequenc
c 680	34	2.0	524	22	AAK73289	Human immune/haema	c 753	34	2.0	2251	24	AAAL5101	Human fascicular a
c 681	34	2.0	526	24	ABN60140	Human cancer relat	c 754	34	2.0	2272	22	ABA07732	Human ovarian and
c 682	34	2.0	528	22	AAK65355	Human immune/haema	c 755	34	2.0	2272	22	AAAL02963	Human reproductive
c 683	34	2.0	532	22	AAH70185	Human cervical can	c 756	34	2.0	2425	24	ABL41330	Human mitochondria
c 684	34	2.0	539	22	AAAL15376	Human breast cance	c 757	34	2.0	2571	22	AAAL35850	Human musculoskele
c 685	34	2.0	544	22	AAF94188	Primer specific fo	c 758	34	2.0	2571	22	AAAL35851	Human musculoskele
c 686	34	2.0	547	23	ABV61418	Human prostate exp	c 759	34	2.0	2761	22	AAAL35852	Human musculoskele
c 687	34	2.0	554	24	ABN62717	Human cancer relat	c 760	34	2.0	2761	22	AAAL07034	Human reproductive
c 688	34	2.0	555	24	ABQ56590	Human colon cancer	c 761	34	2.0	2761	22	AAAL07034	Human reproductive
c 689	34	2.0	568	22	AAH09350	Human cDNA clone (	c 762	34	2.0	2761	22	AAAL07034	Human reproductive
c 690	34	2.0	579	24	ABN65004	Human cancer relat	c 763	34	2.0	2761	22	AAAL07034	Human reproductive
c 691	34	2.0	588	22	AAH09328	Human cDNA clone (	c 764	34	2.0	2761	22	AAAL07034	Human reproductive
c 692	34	2.0	589	23	ABV51491	Human prostate exp	c 765	34	2.0	2761	22	AAAL07034	Human reproductive
c 693	34	2.0	590	22	AAH68756	Human cervical can	c 766	34	2.0	2761	22	AAAL07034	Human reproductive
c 694	34	2.0	595	23	ABV45366	Human prostate exp	c 767	34	2.0	2761	22	AAAL07034	Human reproductive
c 695	34	2.0	596	23	ABV39318	Human prostate exp	c 768	34	2.0	2761	22	AAAL07034	Human reproductive
c 696	34	2.0	596	24	ABN61947	Human cancer relat	c 769	34	2.0	2761	22	AAAL07034	Human reproductive
c 697	34	2.0	597	22	AAH70945	Human cancer relat	c 770	34	2.0	2761	22	AAAL07034	Human reproductive
c 698	34	2.0	606	24	ABN65924	Human cervical can	c 771	34	2.0	2761	22	AAAL07034	Human reproductive
c 699	34	2.0	612	23	ABV29554	Human prostate exp	c 772	34	2.0	2761	22	AAAL07034	Human reproductive
c 700	34	2.0	617	23	ABV50143	Human prostate exp	c 773	34	2.0	2761	22	AAAL07034	Human reproductive
c 701	34	2.0	619	22	AAH65181	Novel human polynu	c 774	34	2.0	2761	22	AAAL07034	Human reproductive
c 702	34	2.0	626	22	AAH70182	Human cervical can	c 775	34	2.0	2761	22	AAAL07034	Human reproductive
c 703	34	2.0	629	22	AAH70158	Human cervical can	c 776	34	2.0	2761	22	AAAL07034	Human reproductive
c 704	34	2.0	634	22	AAH70184	Human cervical can	c 777	34	2.0	2761	22	AAAL07034	Human reproductive
c 705	34	2.0	638	23	ABV44845	Human prostate exp	c 778	34	2.0	2761	22	AAAL07034	Human reproductive
c 706	34	2.0	642	24	ABN63166	Human cancer relat	c 779	34	2.0	2761	22	AAAL07034	Human reproductive
c 707	34	2.0	646	22	AAAS34041	Human cDNA encodin	c 780	34	2.0	2761	22	AAAL07034	Human reproductive
c 708	34	2.0	655	21	AAAL24222	Human breast cance	c 781	34	2.0	2761	22	AAAL07034	Human reproductive
c 709	34	2.0	676	21	AAAL7674	Human secreted pro	c 782	34	2.0	2761	22	AAAL07034	Human reproductive
c 710	34	2.0	695	21	AAAL7173	N. meningitidis pa	c 783	34	2.0	2761	22	AAAL07034	Human reproductive
c 711	34	2.0	700	22	AAH91990	Human inflammatory	c 784	34	2.0	2761	22	AAAL07034	Human reproductive
c 712	34	2.0	700	22	AAH91991	Human inflammatory	c 785	34	2.0	2761	22	AAAL07034	Human reproductive
c 713	34	2.0	700	22	AAH92983	Human inflammatory	c 786	34	2.0	2761	22	AAAL07034	Human reproductive
c 714	34	2.0	700	22	AAH92984	Human inflammatory	c 787	34	2.0	2761	22	AAAL07034	Human reproductive
c 715	34	2.0	735	23	ABV20812	Human prostate exp	c 788	34	2.0	2761	22	AAAL07034	Human reproductive
c 716	34	2.0	735	23	ABV21792	Human prostate exp	c 789	34	2.0	2761	22	AAAL07034	Human reproductive
c 717	34	2.0	735	23	ABV26659	Human prostate exp	c 790	34	2.0	2761	22	AAAL07034	Human reproductive
c 718	34	2.0	735	23	ABV27615	Human prostate exp	c 791	34	2.0	2761	22	AAAL07034	Human reproductive
c 719	34	2.0	752	22	AAH07160	Human cDNA clone (	c 792	34	2.0	2761	22	AAAL07034	Human reproductive
c 720	34	2.0	847	24	ABQ89483	Human prostate exp	c 793	34	2.0	2761	22	AAAL07034	Human reproductive
c 721	34	2.0	857	22	AAH03643	Human cDNA clone (	c 794	34	2.0	2761	22	AAAL07034	Human reproductive
c 722	34	2.0	870	23	AAAS0091	Human cDNA clone (	c 795	34	2.0	2761	22	AAAL07034	Human reproductive
c 723	34	2.0	1022	22	AAAS36702	DNA encoding novel	c 796	34	2.0	2761	22	AAAL07034	Human reproductive
c 724	34	2.0	1022	22	AAAS36709	Human cardiovascular	c 797	34	2.0	2761	22	AAAL07034	Human reproductive
c 725	34	2.0	1042	23	ABK42533	Human cardiovascular	c 798	34	2.0	2761	22	AAAL07034	Human reproductive
c 726	34	2.0	1094	20	AAV84442	Genomic sequence #	c 799	34	2.0	2761	22	AAAL07034	Human reproductive
c 727	34	2.0	1094	22	ABAB3225	Human secreted pro	c 800	34	2.0	2761	22	AAAL07034	Human reproductive
c 728	34	2.0	1160	20	AAK61434	DNA encoding a hum	c 801	34	2.0	2761	22	AAAL07034	Human reproductive
c 729	34	2.0	1162	22	AAK85130	Human immune/haema	c 802	34	2.0	2761	22	AAAL07034	Human reproductive
c 730	34	2.0	1184	22	AAK85128	Human immune/haema	c 803	34	2.0	2761	22	AAAL07034	Human reproductive
c 731	34	2.0	1213	22	AAK85129	Human immune/haema	c 804	34	2.0	2761	22	AAAL07034	Human reproductive
c 732	34	2.0	1268	22	AAK71636	Human immune/haema	c 805	34	2.0	2761	22	AAAL07034	Human reproductive
c 733	34	2.0	1268	22	AAK71636	Human immune/haema	c 806	34	2.0	2761	22	AAAL07034	Human reproductive
c 734	34	2.0	1268	22	AAK71636	Human immune/haema	c 807	34	2.0	2761	22	AAAL07034	Human reproductive
c 735	34	2.0	1275	20	AAV84529	Human musculoskele	c 808	34	2.0	2761	22	AAAL07034	Human reproductive
c 736	34	2.0	1275	20	AAV84529	Human musculoskele	c 809	34	2.0	2761	22	AAAL07034	Human reproductive
c 737	34	2.0	1275	20	AAV84529	Human secreted pro	c 810	34	2.0	2761	22	AAAL07034	Human reproductive
c 738	34	2.0	1393	22	ABAI83312	Human secreted pro	c 811	34	2.0	2761	22	AAAL07034	Human reproductive
c 739	34	2.0	1438	22	AAH15361	Human nervous syst	c 812	34	2.0	2761	22	AAAL07034	Human reproductive
c 739	34	2.0	1497	22	AAK87499	Human cDNA sequenc	c 813	34	2.0	2761	22	AAAL07034	Human reproductive
c 739	34	2.0	1497	22	AAK87499	Human immune/haema	c 814	34	2.0	2761	22	AAAL07034	Human reproductive

c 813	34	2.0	11185	22	AAS28662	Genomic sequence #	c 886	34	2.0	22452	22	AAS27827	DNA encoding novel
c 814	34	2.0	11464	19	AAV48228	Interleukin 18 con	c 887	34	2.0	22452	22	AAS27829	DNA encoding novel
c 815	34	2.0	11599	22	ABA18087	Human nervous syst	c 888	34	2.0	22713	22	AAK64950	Human immune/haema
c 816	34	2.0	11789	22	AAK65081	Human immune/haema	c 889	34	2.0	23822	24	ABL67538	Thyroid cancer rel
c 817	34	2.0	11839	22	AAAL04940	Human reproductive	c 890	34	2.0	23996	22	ABA18618	Human nervous syst
c 818	34	2.0	11839	23	ABL97834	Human testicular a	c 891	34	2.0	24279	22	AAK74999	Human immune/haema
c 819	34	2.0	11990	24	ABL61824	Colon adenocarcino	c 892	34	2.0	24329	22	AAK75000	Human immune/haema
c 820	34	2.0	11990	24	ABL69703	Prostate cancer re	c 893	34	2.0	24329	22	AAK75856	Human immune/haema
c 821	34	2.0	11992	22	AAI99007	Human excretory re	c 894	34	2.0	24908	22	AAK81665	Human immune/haema
c 822	34	2.0	11992	22	AAK80050	Human immune/haema	c 895	34	2.0	26110	22	AAK65036	Human immune/haema
c 823	34	2.0	11992	22	AAI63357	Human kidney relat	c 896	34	2.0	26110	22	AAK78526	Human immune/haema
c 824	34	2.0	12500	22	AAK77971	Human immune/haema	c 897	34	2.0	27154	22	AAAL03708	Human reproductive
c 825	34	2.0	12555	22	ABA18329	Human nervous syst	c 898	34	2.0	27675	22	AAK85837	Human immune/haema
c 826	34	2.0	13158	22	AAK68742	Human immune/haema	c 899	34	2.0	27681	22	AAK36497	Human immune/haema
c 827	34	2.0	13255	22	AAK76842	Human immune/haema	c 900	34	2.0	27681	22	AAK36498	Human cardiovascular
c 828	34	2.0	13310	22	AAK75442	Human immune/haema	c 901	34	2.0	27681	22	AAK85843	Human immune/haema
c 829	34	2.0	13310	22	AAK89617	Human digestive sy	c 902	34	2.0	27754	24	ABQ72998	Human transporter
c 830	34	2.0	13409	22	ABA08135	Human ovarian and	c 903	34	2.0	28149	22	ABA17961	Human nervous syst
c 831	34	2.0	13409	22	AAAL08913	Human reproductive	c 904	34	2.0	28588	22	AAK89418	Human digestive sy
c 832	34	2.0	14282	22	AAK68418	Human immune/haema	c 905	34	2.0	28588	22	AAK31923	Human liver associ
c 833	34	2.0	14543	24	ABK15798	Human von Hippel-L	c 906	34	2.0	28588	24	ABN90278	Human liver antige
c 834	34	2.0	14777	22	AAK84503	Human immune/haema	c 907	34	2.0	28994	19	AAV15826	Genomic DNA for in
c 835	34	2.0	14925	22	AAAL03282	Human reproductive	c 908	34	2.0	30393	22	AAK67239	Human immune/haema
c 836	34	2.0	14925	22	AAAL04552	Human reproductive	c 909	34	2.0	31348	22	AAK35759	Human cardiovascular
c 837	34	2.0	14925	23	ABL97475	Human testicular a	c 910	34	2.0	31348	22	AAK79227	Human immune/haema
c 838	34	2.0	15037	22	AAI99283	Human excretory re	c 911	34	2.0	31741	22	ABA21231	Human nervous syst
c 839	34	2.0	15037	22	AAI63633	Human kidney relat	c 912	34	2.0	31994	22	AAK30619	DNA encoding novel
c 840	34	2.0	15196	22	AAK73103	Human immune/haema	c 913	34	2.0	31994	22	AAK28165	Genomic sequence #
c 841	34	2.0	15196	22	AAK87548	Human immune/haema	c 914	34	2.0	32012	22	AAAL05864	Human reproductive
c 842	34	2.0	15266	22	AAK37432	Human musculoskele	c 915	34	2.0	32012	23	ABL98428	Human testicular a
c 843	34	2.0	15266	22	AAK73549	Human immune/haema	c 916	34	2.0	32152	22	AAK39621	Genomic sequence #
c 844	34	2.0	15271	22	AAAL37433	Human musculoskele	c 917	34	2.0	32152	22	AAK89020	Human digestive sy
c 845	34	2.0	15271	22	AAK73550	Human immune/haema	c 918	34	2.0	32152	22	AAK91534	Human digestive sy
c 846	34	2.0	15484	22	AAK70808	Human immune/haema	c 919	34	2.0	32152	22	AAI57791	Human colorectal c
c 847	34	2.0	15682	22	AAK89506	Human digestive sy	c 920	34	2.0	32169	22	ABA14358	Human nervous syst
c 848	34	2.0	16100	22	AAAL37333	Human musculoskele	c 921	34	2.0	32191	22	ABA07814	Human ovarian and
c 849	34	2.0	17216	22	ABA18237	Human nervous syst	c 922	34	2.0	32191	22	AAAL03616	Human reproductive
c 850	34	2.0	17216	22	AAAL37200	Human musculoskele	c 923	34	2.0	32192	22	ABA17963	Human nervous syst
c 851	34	2.0	17217	22	ABA18239	Human nervous syst	c 924	34	2.0	32194	22	AAAL04340	Human reproductive
c 852	34	2.0	17217	22	AAAL37201	Human musculoskele	c 925	34	2.0	32204	22	AAK39620	Genomic sequence #
c 853	34	2.0	17245	22	AAK83897	Human immune/haema	c 926	34	2.0	32204	22	AAK89019	Human digestive sy
c 854	34	2.0	17357	22	AAK86680	Human immune/haema	c 927	34	2.0	32204	22	AAK91533	Human digestive sy
c 855	34	2.0	17363	22	AAK86681	Human immune/haema	c 928	34	2.0	32204	22	AAI57790	Human colorectal c
c 856	34	2.0	17481	22	AAK68583	Human immune/haema	c 929	34	2.0	32205	22	ABA08217	Human ovarian and
c 857	34	2.0	17481	22	AAK73255	Human immune/haema	c 930	34	2.0	32205	22	AAAL07525	Human reproductive
c 858	34	2.0	17481	22	AAK75443	Human immune/haema	c 931	34	2.0	32353	22	AAK70003	Human immune/haema
c 859	34	2.0	17481	22	AAK89618	Human digestive sy	c 932	34	2.0	34831	22	AAK82518	Human immune/haema
c 860	34	2.0	17590	21	AAZ50904	Human TBC-1 partia	c 933	34	2.0	34917	22	AAK70686	Human immune/haema
c 861	34	2.0	17949	22	AAK89228	Human signal trans	c 934	34	2.0	36305	24	ABK22783	Human high bone ma
c 862	34	2.0	18188	22	AAK87552	Human immune/haema	c 935	34	2.0	36933	22	AAK66362	Human immune/haema
c 863	34	2.0	18252	22	AAK87554	Human immune/haema	c 936	34	2.0	37314	22	AAK71358	Human immune/haema
c 864	34	2.0	18389	22	AAK74864	Human immune/haema	c 937	34	2.0	37783	22	AAK70780	Human immune/haema
c 865	34	2.0	18389	22	AAK82632	Human immune/haema	c 938	34	2.0	37783	22	AAK76625	Human immune/haema
c 866	34	2.0	18402	22	AAK33381	DNA encoding human	c 939	34	2.0	37783	22	AAK80913	Human immune/haema
c 867	34	2.0	18403	22	AAK33382	DNA encoding human	c 940	34	2.0	41684	21	AAK28150	Human immune/haema
c 868	34	2.0	18533	22	AAK75941	Human immune/haema	c 941	34	2.0	42324	22	AAK84724	Human immune/haema
c 869	34	2.0	18647	21	AAF21059	Human low adenosin	c 942	34	2.0	44211	22	AAK85974	Human immune/haema
c 870	34	2.0	18648	21	AAA34937	Human adenosine re	c 943	34	2.0	45186	22	AAE60478	Wild-type human CT
c 871	34	2.0	18648	24	ABL65840	Lung cancer relate	c 944	34	2.0	48000	22	AAK72996	Human calcium sens
c 872	34	2.0	18648	24	ABL65840	Lung cancer relate	c 945	34	2.0	48203	22	AAK70161	Human immune/haema
c 873	34	2.0	18925	22	ABA81553	Kidney cancer rela	c 946	34	2.0	48203	22	AAK81663	Human immune/haema
c 874	34	2.0	18925	24	AAK94565	Human phospholipid	c 947	34	2.0	48203	22	AAK83628	Human immune/haema
c 875	34	2.0	18925	24	AAK94565	Human phospholipid	c 948	34	2.0	48204	22	AAK70164	Human immune/haema
c 876	34	2.0	19408	22	AAK42003	Genomic sequence #	c 949	34	2.0	48204	22	AAK81666	Human immune/haema
c 877	34	2.0	19408	22	AAK87230	Human immune/haema	c 950	34	2.0	48204	22	AAK82630	Human immune/haema
c 878	34	2.0	19408	22	AAK90644	Human digestive sy	c 951	34	2.0	50849	24	ABN87883	Human glutathione
c 879	34	2.0	19596	22	AAK73967	Human immune/haema	c 952	34	2.0	53552	22	AAK31655	Genomic DNA sequen
c 880	34	2.0	19596	22	AAK73967	Human immune/haema	c 953	34	2.0	54108	24	ABK22782	Human high bone ma
c 881	34	2.0	19929	22	AAAL07279	Human reproductive	c 954	34	2.0	57273	24	ABK22782	Human high bone ma
c 882	34	2.0	19929	23	ABL98825	Human testicular a	c 955	34	2.0	59065	24	ABL42416	Human serine/threo
c 883	34	2.0	20507	22	AAAL07278	Human reproductive	c 956	34	2.0	66933	22	ABA82625	Human HBM gene reg
c 884	34	2.0	20907	23	ABL98824	Human testicular a	c 957	34	2.0	72049	22	ABA82623	Human HBM gene reg
c 885	34	2.0	21696	22	AAK85713	Human immune/haema	c 958	34	2.0	81800	24	ABK84756	Human cDNA differe



Db	301		GCATCATTCACAATGAGGGGAGCCACAGCCTGGGCAACAGAGAAATCTCCGTC	360
Qy	361	CAAAAAAAAAAAAAAAAAAGRATTAGGCTGGTGGCTGTAGTCCAGCTACTT	420	
Db	361	CAAAAAAAAAAAAAAAAAAGRATTAGGCTGGTGGCTGTAGTCCAGCTACTT	420	
Qy	421	GGGAGCAGGGGTCCACTTGATGTCGAGACTGCAGTGAAGCCATGATCCTGCCACTGCAC	480	
Db	421	GGGAGCAGGGGTCCACTTGATGTCGAGACTGCAGTGAAGCCATGATCCTGCCACTGCAC	480	
Qy	481	TCCGGCTGGGCACACAGAGTGAGACCTGTCTTAAGAAAAAAAAAATAAGCACATATC	540	
Db	481	TCCGGCTGGGCACACAGAGTGAGACCTGTCTTAAGAAAAAAAAAATAAGCACATATC	540	
Qy	541	CTGAACAAGAGTCCTCCATAAGCTTTCCACCACAGATTTCTAATCAGAAACATGGAGGCCA	600	
Db	541	CTGAACAAGAGTCCTCCATAAGCTTTCCACCACAGATTTCTAATCAGAAACATGGAGGCCA	600	
Qy	601	GAAGCAGTGGAGAGACRACCTCAGGACGCCGGGAGGATGTTGTACAGGCTGGGG	660	
Db	601	GAAGCAGTGGAGAGACRACCTCAGGACGCCGGGAGGATGTTGTACAGGCTGGGG	660	
Qy	661	CAAGGCCCTTCCGGCTACCACTGGGAGCTCTGGGACAGCCCTGTTGCAACAAAGAAGC	720	
Db	661	CAAGGCCCTTCCGGCTACCACTGGGAGCTCTGGGACAGCCCTGTTGCAACAAAGAAGC	720	
Qy	721	CATAGCCCGGCACAGACCCAGGAATGTGGCTGGCTGGGACAGCCCTGTTGCAACAAAGAAGC	780	
Db	721	CATAGCCCGGCACAGACCCAGGAATGTGGCTGGGCTGGGACAGCCCTGTTGCAACAGAGT	780	
Qy	781	GGTCCCATCCAGGAAACCTCCGGCATGGCTGGGAAAGTGGGTACTTGGTGGCGGTCTGT	840	
Db	781	GGTCCCATCCAGGAAACCTCCGGCATGGCTGGGAAAGTGGGTACTTGGTGGCGGTCTGT	840	
Qy	841	ATGTGTGTGTACTGGTGTGTGACAGAGAATGTGCTCTAAGTGTCAAGTGTGAGTGT	900	
Db	841	ATGTGTGTGTACTGGTGTGTGACAGAGAATGTGCTCTAAGTGTCAAGTGTGAGTGT	900	
Qy	901	GTGTATGTGTGAATATTTGTTCTTGTGGTGCATTTTCTGCTGTGTAATCGTGTCCCTG	960	
Db	901	GTGTATGTGTGAATATTTGTTCTTGTGGGTGATTTTCTGCTGTGTAATCGTGTCCCTG	960	
Qy	961	CAAGTGTGAACAAGTGGACAAGTGTCTGGAGTGGACAAGAGATCTGTGCACCATCAGT	1020	
Db	961	CAAGTGTGAACAAGTGGACAAGTGTCTGGAGTGGACAAGAGATCTGTGCACCATCAGT	1020	
Qy	1021	GTGTGCATACGCTGTGTCATGTCAAGAGTGCAGGTGAAGTGAAGGACACAGGCCCATG	1080	
Db	1021	GTGTGCATACGCTGTGTCATGTCAAGAGTGCAGGTGAAGTGAAGGACACAGGCCCATG	1080	
Qy	1081	ATGCCACTCATATCAGGAGCTTAAGGCCCCAGAGTACTGCCAGTGACAGATAGGGTG	1140	
Db	1081	ATGCCACTCATATCAGGAGCTTAAGGCCCCAGAGTACTGCCAGTGACAGATAGGGTG	1140	
Qy	1141	CTGAAGTCACTCTGGAGTGGGAGTGGGGTATGGGAAAGGCAAGGCCATGTTCTTGA	1200	
Db	1141	CTGAAGTCACTCTGGAGTGGGAGTGGGGTATGGGAAAGGCAAGGCCATGTTCTTGA	1200	
Qy	1201	GGAGGGTGTGACTACATTAGGGTGTATGAGCCTAGCTGGGAGGTGGATGCCCGGTCC	1260	
Db	1201	GGAGGGTGTGACTACATTAGGGTGTATGAGCCTAGCTGGGAGGTGGATGCCCGGTCC	1260	
Qy	1261	ACTGAACCCCTGTTATCCCAAGAGCTTTGAGGCTTTAGGAGCTGGAGTGGGGAGAG	1320	
Db	1261	ACTGAACCCCTGTTATCCCAAGAGCTTTGAGGCTTTAGGAGCTGGAGTGGGGAGAG	1320	
Qy	1321	GGGTGACTTCTCCGACACAGGCCCTCCACCGGCTACCCCTGGGTAAAGGGCTGAGCAG	1380	
Db	1321	GGGTGACTTCTCCGACACAGGCCCTCCACCGGCTACCCCTGGGTAAAGGGCTGAGCAG	1380	
Qy	1381	GAAGCAGGGCAAGAACCTCTGGAGAGCCCATACCCGCCCTGGCCTGACTCTGCCACTG	1440	

Db	1381	GAAGCAGGGGCAAGAACCTCTGGAGCAGAGCCCATACCCGGCCCTGGCCTGACTCTGCCACTG	1414
Qy	1441	GCACACAGTCAACACACAGCAGGTTCACCTCACAGCAGAGGGCAAGGCCATCATCAGCTCC	1500
Db	1441	GCACACAGTCAACACACAGCAGGTTCACCTCACAGCAGAGGGCAAGGCCATCATCAGCTCC	1500
Qy	1501	CTTTATAAGGGAAGGGTCACGCCTCGGTGCTGCTGAGAGTGCTCTGCTGCTGCTCTGCTG	1560
Db	1501	CTTTATAAGGGAAGGGTCACGCCTCGGTGCTGCTGAGAGTGCTCTGCTGCTCTGCTGCTG	1560
Qy	1561	CCTGCTGGGTGGGGTGCACAGTGTGTCACAGAGGAGCCCATTTGGTAGTGAGGACAGTA	1620
Db	1561	CCTGCTGGGTGGGGTGCACAGTGTGTCACAGAGGAGCCCATTTGGTAGTGAGGACAGTA	1620
Qy	1621	TGGGGCTAGAAGCACTGGTCCCTGGCCGTGATAGTGCCCATCTTCCTGCTCTCTGCTGG	1680
Db	1621	TGGGGCTAGAAGCACTGGTCCCTGGCCGTGATAGTGCCCATCTTCCTGCTCTCTGCTGG	1680
RESULT 2			
ID	AAH26169		
XX	AAH26169 standard; DNA; 1669 BP.		
AC	AAH26169;		
DT	17-SEP-2001 (first entry)		
XX	Human cytochrome P450 CYP2D6 gene promoter region.		
DE	Cytochrome P450; CYP2D6; promoter; drug metabolism; human;		
KW	diagnosis; therapy; ds.		
KW	Homo sapiens.		
OS	Homo sapiens.		
XX			
PH	Key		
FT	primer_bind	Location/Qualifiers	
FT		complement (14..36)	
FT		/*tag= a	
FT	primer_bind	/*note= "amplification primer upf14"	
FT		337..358	
FT		/*tag= b	
FT	primer_bind	/*note= "sequence primer R1"	
FT		493..514	
FT		/*tag= c	
FT	primer_bind	/*note= "sequencing primer R2"	
FT		complement (565..577)	
FT		/*tag= d	
FT	primer_bind	/*note= "sequencing primer F2"	
FT		602..620	
FT		/*tag= e	
FT	primer_bind	/*note= "sequencing primer R3"	
FT		complement (968..988)	
FT		/*tag= f	
FT	primer_bind	/*note= "sequencing primer F3"	
FT		1124..1143	
FT		/*tag= g	
FT	primer_bind	/*note= "sequencing primer R4"	
FT		1605..1623	
FT		/*tag= h	
FT	primer_bind	/*note= "sequencing primer R5"	
FT		1650..1669	
FT		/*tag= i	
FT	5'UTR	/*note= "amplification primer upr1669"	
FT		1532..1619	
FT		/*tag= j	
FT	CDS	1620..1669	
FT		/*tag= k	
FT		/partial	
FT		/*note= "5' region of CYP2D6 coding region"	
XX			
PN	WO200155432-A2.		
XX			
PD	02-AUG-2001.		
XX			

30-JAN-2001; 2001WO-EP00954.  
31-JAN-2000; 2000EP-0101889.  
(EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
Raimundo S, Zanger U;  
WPI; 2001-457734/49.  
A polynucleotide capable of hybridizing to CYP2D6 promoter useful for  
the optimization of drug therapies using substrates of cytochrome P-450  
-  
Claim 1; Fig 1; 41pp; English.  
The present sequence is that of the promoter region of the human  
cytochrome P450 CYP2D6 gene. The promoter region was amplified  
by PCR from leukocyte DNA of over 50 individuals, and sequenced.  
8 previously unknown single nucleotide polymorphisms (SNP) were  
identified. These were at: base 36 (base -1584 according to the  
Human Cytochrome P450 Allele Nomenclature), where the SNP was C to  
G, occurring at an estimated frequency of approximately 20% in the  
whole population, and resulting in increased enzyme activity;  
position 194 (-1426), C to T, approximately 20% frequency; neutral  
function; position 385 (-1235), A to G, approximately 50% frequency,  
neutral function; position 620 (-1000), G to A, approximately 20%  
frequency, neutral function; position 880 (-740), C to T,  
approximately 30% frequency, unknown function; position 940 (-680),  
G to A, approximately 30% frequency, unknown function; 1255 (-365),  
G to A, rare, unknown function; and 1298 (-322), T to C, rare,  
unknown function. The C to G mutation at -1584 bp is strongly  
associated with lower metabolic ratios, and a molecular variant  
polynucleotide having G at this position is claimed (see AAH26179).  
The invention provides a method of diagnosing a disorder related to  
reduced or enhanced capacity for clearance of CYP2D6 substrates  
(antiarrhythmic, beta-adrenergic receptor-antagonist, tricyclic  
antidepressant, selective serotonin reuptake inhibitor, neuroleptic,  
opioid, cytosolic or amphetamine), or susceptibility to such a  
disorder, by determining the presence of a mutation in the CYP2D6  
promoter. The strong association of the common C to G mutation at  
-1584 bp with increased enzyme activity significantly improves the  
correlation between genotype and phenotype in the CYP2D6 polymorphism.  
Testing for the mutation will allow the identification of intermediate  
metabolizers and therefore allow quantitative predictions to be made  
on in vivo drug metabolism capacity, thus providing a very potent  
tool for improving the therapy of diseases with drugs that are  
targets of the CYP2D6 gene product.  
Sequence 1669 BP; 413 A; 376 C; 534 G; 338 T; 8 other;  
Query Match 90.2%; Score 1516; DB 22; Length 1669;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1666; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GAATCAAGACCAAGCTGGACAACTTGGAGAACCCGGTCTCTACAAAAAATACAAAAT 60  
Db 1 GAATCAAGACCAAGCTGGACAACTTGGAGAACCCGGTCTCTACAAAAAATACAAAAT 60  
QY 61 AGCTGGGATTGGGTGGGTGGCTGCATGCCTATATATCCAGACCTTTGGGAGCCTGAGGTG 120  
Db 61 AGCTGGGATTGGGTGGGTGGCTGCATGCCTATATATCCAGACCTTTGGGAGCCTGAGGTG 120  
QY 121 GGTGGATACCTGAATCAGGAGTTCAGACTAGCTAGCTGGCCCAACATGGTGAACCCCTATC 180  
Db 121 GGTGGATACCTGAATCAGGAGTTCAGACTAGCTAGCTGGCCCAACATGGTGAACCCCTATC 180  
QY 181 TCTACTGAAAATAYAAAAAGCTAGAGCTGGTGGCACACACCTGTAATCCACGCTACTTAG 240  
Db 181 TCTACTGAAAATAYAAAAAGCTAGAGCTGGTGGCACACACCTGTAATCCACGCTACTTAG 240  
QY 241 GAGGCTGAGCGAGGAGAAATGGTTGAAGCCTAGAGGTGAGGTGAGTGCAGCCAGATT 300  
Cc , ,  
Cc , ,

Db 241 GAGGCTGAGCGAGGAGAAATTCCTTGAAGCCTAGAGGTGAAGTTGTAGTGAGCCGAGATT 300  
QY 301 GCATCATTTGCACATGAGGGGAGCCACACAGCCCTGGGCAACAGAGGAAATCTCCGTCTC 360  
Db 301 GCATCATTTGCACATGAGGGGAGCCACACAGCCCTGGGCAACAGAGGAAATCTCCGTCTC 360  
QY 361 CAAAAAAGGATCCCTCCATTAAGCTGGGTGGTGCCTGTAGTCCCTAGCTACTT 420  
Db 361 CAAAAAAGGATCCCTCCATTAAGCTGGGTGGTGCCTGTAGTCCCTAGCTACTT 420  
QY 421 GGGAGGAGGGGTCCACTTGAATGTCGAGACTGCAGTGAAGCCATGATCCCTAGCTGCAC 480  
Db 421 GGGAGGAGGGGTCCACTTGAATGTCGAGACTGCAGTGAAGCCATGATCCCTAGCTGCAC 480  
QY 481 TCCGGCTGGGCAACAGAGTGAGACCTGTCTAAAGAAAAAATAAAGCAACATATC 540  
Db 481 TCCGGCTGGGCAACAGAGTGAGACCTGTCTAAAGAAAAAATAAAGCAACATATC 540  
QY 541 CTGAACAAAGGATCCCTCCATTAAGCTGGGTGGTGCCTGTAGTCCCTAGCTGCAC 600  
Db 541 CTGAACAAAGGATCCCTCCATTAAGCTGGGTGGTGCCTGTAGTCCCTAGCTGCAC 600  
QY 601 GAAAGCAGTGGAGGAGACRACCTCAGGCAGCCCGGAGGATGTTGTACAGGCTGGG 660  
Db 601 GAAAGCAGTGGAGGAGACRACCTCAGGCAGCCCGGAGGATGTTGTACAGGCTGGG 660  
QY 661 CAAGGCTCCCGCTTACCAACTGGGAGCTCTGGGAGCAGCCCTGTGCAACAAAGC 720  
Db 661 CAAGGCTCCCGCTTACCAACTGGGAGCTCTGGGAGCAGCCCTGTGCAACAAAGC 720  
QY 721 CATAGCCCGGCCAGAGCCAGGAATGTGGCTGGGAGCAGCCTCTGGACAGGAGT 780  
Db 721 CATAGCCCGGCCAGAGCCAGGAATGTGGCTGGGAGCAGCCTCTGGACAGGAGT 780  
QY 781 GGTCCATCCAGGAAACCTCCGCATGGCTGGGAGTGGGATGCTGTGCGGGTCTGT 840  
Db 781 GGTCCATCCAGGAAACCTCCGCATGGCTGGGAGTGGGATGCTGTGCGGGTCTGT 840  
QY 841 ATGTGTGTGACTGTGTGTGAGAGAGATGTGCTGCTAAGTGTCAGTGTGAGTCT 900  
Db 841 ATGTGTGTGACTGTGTGTGAGAGAGATGTGCTGCTAAGTGTCAGTGTGAGTCT 900  
QY 901 GTGTATGTGTAATATGTCTTGTGTGGTGATTTTCTGCTGTGTAATCTGCTGCTG 960  
Db 901 GTGTATGTGTAATATGTCTTGTGTGGTGATTTTCTGCTGTGTAATCTGCTGCTG 960  
QY 961 CAAAGTGAACAAAGTGGACAAAGTGTCTGGGAGTGGGAGATCTGTGCACCATCAGT 1020  
Db 961 CAAAGTGAACAAAGTGGACAAAGTGTCTGGGAGTGGGAGATCTGTGCACCATCAGT 1020  
QY 1021 GTGTGCATAGCCTCTGTGCATGTCAAGAGTGAAGTGAAGGAGCCAGGCCCATG 1080  
Db 1021 GTGTGCATAGCCTCTGTGCATGTCAAGAGTGAAGTGAAGGAGCCAGGCCCATG 1080  
QY 1081 ATGCCACTCATCATCAGGAGCTTAAAGCCCAAGTGAAGTGAAGGAGCCAGGCCCATG 1140  
Db 1081 ATGCCACTCATCATCAGGAGCTTAAAGCCCAAGTGAAGTGAAGGAGCCAGGCCCATG 1140  
QY 1141 CTGAAGGTCACCTCTGGAGTGGCAGTGGGAGTGGGAGGAGGAGGAGGAGGAGGAG 1200  
Db 1141 CTGAAGGTCACCTCTGGAGTGGCAGTGGGAGTGGGAGGAGGAGGAGGAGGAGGAG 1200  
QY 1201 GGAGGGTTGTGACTACATTTAGGGTGTATGACCTAGCTGGGAGGAGGAGGAGGAGGAG 1260  
Db 1201 GGAGGGTTGTGACTACATTTAGGGTGTATGACCTAGCTGGGAGGAGGAGGAGGAGGAG 1260  
QY 1261 ACTGAAACCCCTGGTATCCCAAGAGGCTTTGAGGCTTTCAGGAGCTTTCAGGAGGAG 1320  
Db 1261 ACTGAAACCCCTGGTATCCCAAGAGGCTTTGAGGCTTTCAGGAGCTTTCAGGAGGAG 1320  
QY 1321 GGGGTGACTTTCGCCAGCAGGCCCTCCACCGGCTACCTCGGTAAAGGCTCGGAGCAG 1380  
Db 1321 GGGGTGACTTTCGCCAGCAGGCCCTCCACCGGCTACCTCGGTAAAGGCTCGGAGCAG 1380





Db 697 ACAGCCCTGTTGTAACAAGAACCATAGCCCGCCAGAGCCCAAGAAATGTGGCTGGGC 756  
QY 757 TGGGAGCAGCCTCTGACAGGAGTGGTCCCATCCAGGAACCTCCGGCATGGCTGGGAAG 816  
Db 757 TGGGAGCAGCCTCTGACAGGAGTGGTCCCATCCAGGAACCTCCGGCATGGCTGGGAAG 816  
QY 817 TGGGGTACTTGGTGCCGGGTCTGTATGTGTGTGACTGCTGTGTGAGAGAAATGTG 876  
Db 817 TGGGGTACTTGGTGCCGGGTCTGTATGTGTGTGACTGCTGTGTGAGAGAAATGTG 876  
QY 877 TGCYTAAGTGTCAAGTGTGAGTCTGTGTATGTGTGTAATATTGTCTTTGTGGGTGATTT 936  
Db 877 TGCYTAAGTGTCAAGTGTGAGTCTGTGTATGTGTGTAATATTGTCTTTGTGGGTGATTT 936  
QY 937 TCTGCRGTGTAAATCGTCTCCCTGCAAGTGTGAACAAGTGGACAAGTCTCTGGAGTGA 996  
Db 937 TCTRCGTGTAAATCGTCTCCCTGCAAGTGTGAACAAGTGGACAAGTCTCTGGAGTGA 996  
QY 997 CAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGCATGTCAAGAGTCAAGG 1056  
Db 997 CAAGAGATCTGTGCACCATCAGGTGTGTGCATAGCGTCTGTGCATGTCAAGAGTCAAGG 1056  
QY 1057 TGAAGTGAAGGACAGGCCCATGATGCCACTCATCAGAGCTCTAAGGCCCCAGGT 1116  
Db 1057 TGAAGTGAAGGACAGGCCCATGATGCCACTCATCAGAGCTCTAAGGCCCCAGGT 1116  
QY 1117 AAGTCCAGTGCACAGATAAGGTGCTGAAGTCACTCTGGAGTGGCAGGTGGGGTAGG 1176  
Db 1117 AAGTCCAGTGCACAGATAAGGTGCTGAAGTCACTCTGGAGTGGCAGGTGGGGTAGG 1176  
QY 1177 GAAAGGCAAGGCCATGTTCTTGGAGGAGGGTGTGTACTACATTAAGGTGTATGAGCCTA 1236  
Db 1177 GAAAGGCAAGGCCATGTTCTTGGAGGAGGGTGTGTACTACATTAAGGTGTATGAGCCTA 1236  
QY 1237 GCTGGAGTGCATGGCCTGCTGACCTGAACCTGCTTATCCAGAGCTTTGACGC 1296  
Db 1237 GCTGGAGTGCATGGCCTGCTGACCTGAACCTGCTTATCCAGAGCTTTGCAAGG 1296  
QY 1297 TTCAGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACAGGCCCTCCACCGGCCT 1356  
Db 1297 TYCAGGAGCTTGGAGTGGGAGAGGGGTGACTTCTCCGACAGGCCCTCCACCGGCCT 1356  
QY 1357 ACCCTGGTGAAGGCCCTTGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1416  
Db 1357 ACCCTGGTGAAGGCCCTTGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1416  
QY 1417 CGCCCTGGCCTGACTCTGCCACTGGCAGCAGTCAACAGCAGAGGTTCACCTCACAGCAG 1476  
Db 1417 CGCCCTGGCCTGACTCTGCCACTGGCAGCAGTCAACAGCAGAGGTTCACCTCACAGCAG 1476  
QY 1477 AGGCAAGGCCATCATCAGCTCCCTTTATAAGGGAAGGTACGCGCTCGGTGTGCTGA 1536  
Db 1477 AGGCAAGGCCATCATCAGCTCCCTTTATAAGGGAAGGTACGCGCTCGGTGTGCTGA 1536  
QY 1537 GAGTGTCTGCCCTGCTCTGCTGCTGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1596  
Db 1537 GAGTGTCTGCCCTGCTCTGCTGCTGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1596  
QY 1597 GCCCATTTGGTAGGCGCAGGTATGGGCTAGAAGCAGTGTGCTGCTGGCCCTGGCCGTGATAG 1656  
Db 1597 GCCCATTTGGTAGGCGCAGGTATGGGCTAGAAGCAGTGTGCTGCTGGCCCTGGCCGTGATAG 1656  
QY 1657 TGGCCATCTTCCT 1669  
Db 1657 TGGCCATCTTCCT 1669

RESULT 4  
AAD34213  
ID AAD34213 standard; DNA; 9432 BP.  
XX  
AC  
AAD34213;  
XX

DT 16-JUL-2002 (first entry)  
XX Human cytochrome P450 2D6 (CYP2D6) gene.  
DE Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
XX ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; ds.  
KW  
XX Homo sapiens.  
OS  
XX WO200218638-A2.  
PN 07-MAR-2002.  
XX  
XX 27-AUG-2001; 2001WO-1B01544.  
PF  
XX 30-AUG-2000; 2000GB-0021286.  
PR  
XX (GEMI-) GEMINI GENOMICS PLC.  
XX  
XX Risinger C, Andersson MK, Lewander T, Oliasson E;  
PI  
XX WPI; 2002-329785/36.  
XX  
XX New sequence determination oligonucleotides, useful for detecting  
PT polymorphic sites in a 5' flanking region of a CYP2D6 gene, as  
PT hybridization probes, as components of diagnostic assays, or in  
PT ligase-based sequence determination -  
XX  
PS Example 3; Fig 1; 63pp; English.  
XX  
XX The invention relates to sequence determination oligonucleotides for  
CC detecting polymorphic sites in a 5' flanking region of cytochrome P450  
CC 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many  
CC different xenobiotics. Human CYP2D6 gene is located on chromosome 22.  
CC The oligonucleotides may be used as in situ hybridisation probes, in  
CC ligase-based sequenced determination, as components of diagnostic assays,  
CC as probes in sequence determination methods based on mismatches, as  
CC hybridisation-based diagnostic assays, and as components of diagnostic  
CC microarray. CYP2D6 is useful to predict variations in an individual's  
CC ability to metabolise certain drugs. The present sequence is human  
CC CYP2D6 gene.  
XX  
SQ Sequence 9432 BP; 1964 A; 2647 C; 2976 G; 1845 T; 0 other;

Query Match 79.6%; Score 1338; DB 24; Length 9432;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1638; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 37 GGTCTCTACAAAAATACAAAATTAGCTGGGATTGGGTGGCTCATGCTATAATC 96  
Db 37 GGTCTCTACAAAAATACAAAATTAGCTGGGATTGGGTGGCTCATGCTATAATC 96  
QY 97 CCAGCACTTTGGAGCCTCAGGTGGGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCC 156  
Db 97 CCAGCACTTTGGAGCCTCAGGTGGGTGGATCACCTGAAGTCAGGAGTTCAAGACTAGCC 156  
QY 157 TGGCCAAACATGTGAACCCCTATCTACTGAAATAYAAAAAGCTAGACGTGGTGGCAC 216  
Db 157 TGGCCAAACATGTGAACCCCTATCTACTGAAATAYAAAAAGCTAGACGTGGTGGCAC 216  
QY 217 ACACCTGTAAATCCAGCTACTTTAGGAGGCTGAGGAGGAGAATTGCTTGAAGCCCTAGAGG 276  
Db 217 ACACCTGTAAATCCAGCTACTTTAGGAGGCTGAGGAGGAGAATTGCTTGAAGCCCTAGAGG 276  
QY 277 TGAAGGTTGTAGTGAGCCGAGATTGCATCATTTGCACAAATGGAGGGAGCCACCGCTGG 336  
Db 277 TGAAGGTTGTAGTGAGCCGAGATTGCATCATTTGCACAAATGGAGGGAGCCACCGCTGG 336  
QY 337 GCAACAAGAGGAAATCTCCGCTCTCCAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 396  
Db 337 GCAACAAGAGGAAATCTCCGCTCTCCAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 396





```
FT FT /*tag= j
FT FT /label= PS8
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution V11M"
FT FT variation
FT FT replace (1100, T)
FT FT /*tag= k
FT FT /label= PS9
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution P34S"
FT FT intron
FT FT 1181..1883
FT FT /*tag= l
FT FT /number= 1
FT FT replace (1827, C)
FT FT /*tag= m
FT FT /label= PS10
FT FT /note= "Novel single nucleotide polymorphism (SNP)"
FT FT variation
FT FT replace (1843, G)
FT FT /*tag= n
FT FT /label= PS11
FT FT /note= "Known single nucleotide polymorphism (SNP)"
FT FT exon
FT FT 1884..2055
FT FT /*tag= o
FT FT /number= 2
FT FT replace (1966, A)
FT FT /*tag= p
FT FT /label= PS12
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution R88H"
FT FT variation
FT FT replace (1974, A)
FT FT /*tag= q
FT FT /label= PS13
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution L91M"
FT FT variation
FT FT replace (1984, G)
FT FT /*tag= r
FT FT /label= PS14
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution H94R"
FT FT variation
FT FT replace (1997, G)
FT FT /*tag= s
FT FT /label= PS15
FT FT /note= "Novel single nucleotide polymorphism (SNP)"
FT FT variation
FT FT replace (2044, C)
FT FT /*tag= t
FT FT /label= PS16
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution V104A"
FT FT variation
FT FT replace (2022, T)
FT FT /*tag= u
FT FT /label= PS17
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT together with PS18 causes the amino acid
FT FT substitution T107F"
FT FT variation
FT FT replace (2023, T)
FT FT /*tag= v
FT FT /label= PS18
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT together with PS17 causes the amino acid
FT FT substitution T107F"
FT FT variation
FT FT replace (2028, G)
FT FT /*tag= w
FT FT /label= PS19
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution I109V"
FT FT variation
FT FT replace (2036, C)
FT FT /*tag= x
FT FT /label= PS20
FT FT /note= "Novel single nucleotide polymorphism (SNP)"
FT FT variation
FT FT replace (2039, T)
FT FT /*tag= y
FT FT /label= PS21
FT FT /note= "Known single nucleotide polymorphism (SNP)"
FT FT intron
FT FT 2056...2605
FT FT /*tag= z
FT FT /number= 2
FT FT /cons_splice= (5'site:NO, 3'site:YES)
FT FT replace (2062, G)
FT FT /*tag= aa
FT FT /label= PS22
FT FT /note= "Novel single nucleotide polymorphism (SNP)"
FT FT variation
FT FT replace (2067, G)
FT FT /*tag= ab
FT FT /label= PS23
FT FT /note= "Novel single nucleotide polymorphism (SNP)"
FT FT variation
FT FT replace (2118, T)
FT FT /*tag= ac
FT FT /label= PS24
FT FT /note= "Novel single nucleotide polymorphism (SNP)"
FT FT variation
FT FT replace (2170, A)
FT FT /*tag= ad
FT FT /label= PS25
FT FT /note= "Known single nucleotide polymorphism (SNP)"
FT FT variation
FT FT replace (2179, C)
FT FT /*tag= ae
FT FT /label= PS26
FT FT /note= "Novel single nucleotide polymorphism (SNP)"
FT FT exon
FT FT 2606...2758
FT FT /*tag= af
FT FT /number= 3
FT FT replace (2611, A)
FT FT /*tag= ag
FT FT /label= PS27
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution F120I"
FT FT variation
FT FT replace (2635, C)
FT FT /*tag= ah
FT FT /label= PS28
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution W128R"
FT FT variation
FT FT replace (2659, A)
FT FT /*tag= ai
FT FT /label= PS29
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT together with PS30 causes the amino acid
FT FT substitution V136I"
FT FT variation
FT FT replace (2661, C)
FT FT /*tag= aj
FT FT /label= PS30
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT together with PS29 causes the amino acid
FT FT substitution V136I"
FT FT variation
FT FT replace (2704, G)
FT FT /*tag= ak
FT FT /label= PS31
FT FT /note= "Known single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution Q151E"
FT FT variation
FT FT replace (2716, A)
FT FT /*tag= al
FT FT /label= PS32
FT FT /note= "Novel single nucleotide polymorphism (SNP);
FT FT causes the amino acid substitution E155R"
FT FT intron
FT FT 2759...2846
FT FT /*tag= am
FT FT /number= 3
FT FT replace (2846, A)
FT FT /*tag= an
FT FT /label= PS33
FT FT /note= "Known single nucleotide polymorphism (SNP)"
FT FT exon
FT FT 2847..3007
FT FT /*tag= ao
FT FT /number= 4
FT FT 3008..3440
FT FT /*tag= ap
FT FT /number= 4
FT FT replace (3292, A)
FT FT variation
```

Query Match 54.0%; Score 907; DB 24; Length 6472;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1057; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 621 ACCCTCAGCAGCCGGAGAGTGTGTACAGGCTGGGGCAGGCGCTTCGGCTACCA 680  
Db 2 ACCCTCAGCAGCCGGAGAGTGTGTACAGGCTGGGGCAGGCGCTTCGGCTACCA 61  
QY 681 ACTGGAGGCTCTGGGAACAGCCCTGTGCAACAAAGAACCCATAGCCCGCCAGAGCCCA 740  
Db 62 ACTGGAGGCTCTGGGAACAGCCCTGTGCAACAAAGAACCCATAGCCCGCCAGAGCCCA 121  
QY 741 GGAATGTGGCTGGGCTGGGAGCAGCCTCTGGACAGAGTGTGTCATCCAGGAACCTC 800  
Db 122 GGAATGTGGCTGGGCTGGGAGCAGCCTCTGGACAGAGTGTGTCATCCAGGAACCTC 181  
QY 801 CGGCATGGCTGGGAAGTGGGGTACTTGGGCGGGTCTATGTGTGTGCTGACTGGTGTG 860  
Db 182 CGGCATGGCTGGGAAGTGGGGTACTTGGGCGGGTCTATGTGTGTGCTGACTGGTGTG 241  
QY 861 TGTGAGAGAGAATGTGTGCCTAAGTGTGAGTGTGAGTGTGTGTGTGTGAATATTGTC 920  
Db 242 TGTGAGAGAGAATGTGTGCCTAAGTGTGAGTGTGTGTGTGTGAATATTGTC 301  
QY 921 TTTGTGTGGGTGATTTTCTGCRGTGTATCTGTGTCCTGCAAGTGTGAACAGTGGACA 980  
Db 302 TTTGTGTGGGTGATTTTCTGCGTGTGTAACTGTGTCCTGCAAGTGTGAACAGTGGACA 361  
QY 981 AGTGTCTGGAGTGACAGAGATCTGTGCACCATCAGGTGTGTCATAGCTCTGTGCA 1040  
Db 362 AGTGTCTGGAGTGACAGAGATCTGTGCACCATCAGGTGTGTCATAGCTCTGTGCA 421  
QY 1041 TGTCAAGAGTGAAGTGAAGTGAAGGACAGGCGCCATGATGCCACTCATCATCAGGAG 1100  
Db 422 TGTCAAGAGTGAAGTGAAGTGAAGGACAGGCGCCATGATGCCACTCATCATCAGGAG 481  
QY 1101 CTCCTAAGGCCCCAGGTGAAGTGCAGTGCAGATAGGGTGTCTGAAGGTCACTCTGGAGTG 1160  
Db 482 CTCCTAAGGCCCCAGGTGAAGTGCAGTGCAGATAGGGTGTCTGAAGGTCACTCTGGAGTG 541  
QY 1161 GGCAGTGGGGTGGGAAAGGAGGCGCATGTTCTGGAGGAGGGTGTGACTACATT 1220  
Db 542 GGCAGTGGGGTGGGAAAGGAGGCGCATGTTCTGGAGGAGGGTGTGACTACATT 601  
QY 1221 AGGGTGTATGAGCTAGCTGGGAGGTGGATGGCCRGCTCCACTGAACCCCTGTTATCCC 1280  
Db 602 AGGGTGTATGAGCTAGCTGGGAGGTGGATGGCCRGCTCCACTGAACCCCTGTTATCCC 661  
QY 1281 AGAAGGCTTTGAGGCTTCAGGAGCTTGGAGTGGGGAGAGGGGTGACTTCTCCGACCAAG 1340  
Db 662 AGAAGGCTTTGAGGCTTCAGGAGCTTGGAGTGGGGAGAGGGGTGACTTCTCCGACCAAG 721  
QY 1341 GCCCTCCACCGGCTTACCCTGGGTAAGGCGCTGGAGCAGGAGGCGGCAAGAACCTC 1400  
Db 722 GCCCTCCACCGGCTTACCCTGGGTAAGGCGCTGGAGCAGGAGGCGGCAAGAACCTC 781  
QY 1401 TGGAGCAGGCCCCATACCGGCGCTGGGCTGACTCTGCCACTGGCAGCACAGTCAACACAGCA 1460  
Db 782 TGGAGCAGGCCCCATACCGGCGCTGGGCTGACTCTGCCACTGGCAGCACAGTCAACACAGCA 841  
QY 1461 GGTTCACCTACAGCAGAGGGCAAGGCCATCATCAGCTCCCTTTTATAAGGGAAGGCTAC 1520  
Db 842 GGTTCACCTACAGCAGAGGGCAAGGCCATCATCAGCTCCCTTTTATAAGGGAAGGCTAC 901  
QY 1521 GCGCTCGGTGCTGAGAGTGTCTGCGCTGGTCCCTGTGCTGTGGGTGGGGTGGCTGCC 1580  
Db 902 GCGCTCGGTGCTGAGAGTGTCTGCGCTGGTCCCTGTGCTGTGGGTGGGGTGGCTGCC 961  
QY 1581 AGGTGTGTCCAGAGAGCCCATTTGTGTAGTGAGGAGGTATGGGGCTAGAAGCACTGGTG 1640  
Db 962 AGGTGTGTCCAGAGAGCCCATTTGTGTAGTGAGGAGGTATGGGGCTAGAAGCACTGGTG 1021  
QY 1641 CCCCTGGCGGTGATAGTGGCCCATCTTCCTGCTCCTGGTGG 1680

Db 1022 CCCCTGGCGGTGATAGTGGCCCATCTTCCTGCTCCTGGTGG 1061  
RESULT 6  
ABQ72215  
ID ABQ72215 standard; DNA; 6472 BP.  
XX  
AC ABQ72215;  
XX  
DT 02-SEP-2002 (first entry)  
XX  
DE Human CYP2D6 gene, SEQ ID NO:1 version #1.  
XX  
KW Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;  
chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;  
antiarrhythmic; arrhythmia; adrenoreceptor antagonist; hypertension;  
tricyclic antidepressant; procainamide; drug induced lupus syndrome;  
environmentally linked disease; Parkinson's disease; haplotyping;  
genotyping; haplotype; genetic variant; single nucleotide polymorphism;  
SNP; drug screening; drug discovery; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT variation replace (636, A)  
FT FT /\*tag= a  
FT FT /label= PS1  
FT FT /note= "Novel single nucleotide polymorphism (SNP);  
given as R in the specification"  
FT FT variation replace (678, C)  
FT FT /\*tag= b  
FT FT /label= PS2  
FT FT /note= "Novel single nucleotide polymorphism (SNP);  
given as Y in the specification"  
FT FT variation replace (769, C)  
FT FT /\*tag= c  
FT FT /label= PS3  
FT FT /note= "Novel single nucleotide polymorphism (SNP);  
given as S in the specification"  
FT FT variation replace (776, G)  
FT FT /\*tag= d  
FT FT /label= PS4  
FT FT /note= "Novel single nucleotide polymorphism (SNP);  
given as R in the specification"  
FT FT variation replace (825, A)  
FT FT /\*tag= e  
FT FT /label= PS5  
FT FT /note= "Known single nucleotide polymorphism (SNP);  
given as R in the specification"  
FT FT variation replace (915, C)  
FT FT /\*tag= f  
FT FT /label= PS6  
FT FT /note= "Novel single nucleotide polymorphism (SNP);  
given as Y in the specification"  
FT CDS 1001..5217  
FT FT /\*tag= g  
FT FT /product= "CYP2D6"  
FT FT exon 1001..1180  
FT FT /\*tag= h  
FT FT /number= 1  
FT FT variation replace (1019, A)  
FT FT /\*tag= i  
FT FT /label= PS7  
FT FT /note= "Known single nucleotide polymorphism (SNP);  
given as R in the specification; causes the  
amino acid substitution V7M"  
FT FT variation replace (1031, A)  
FT FT /\*tag= j  
FT FT /label= PS8  
FT FT /note= "Known single nucleotide polymorphism (SNP);  
given as R in the specification; causes the  
amino acid substitution V11M"  
FT FT

```

FT      variation      replace (1100, T)
FT      /*tag= k
FT      /label= PS9
FT      /note= "Known single nucleotide polymorphism (SNP);
FT      given as Y in the specification; causes the
FT      amino acid substitution P34S"
FT      intron
FT      1181..1883
FT      /*tag= 1
FT      /number= 1
FT      /label= PS10
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as S in the specification"
FT      variation
FT      replace (1843, G)
FT      /*tag= n
FT      /label= PS11
FT      /note= "Known single nucleotide polymorphism (SNP);
FT      given as K in the specification"
FT      exon
FT      1884..2055
FT      /*tag= o
FT      /number= 2
FT      /label= PS12
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as R in the specification; causes the
FT      amino acid substitution R88H"
FT      variation
FT      replace (1974, A)
FT      /*tag= q
FT      /label= PS13
FT      /note= "Known single nucleotide polymorphism (SNP);
FT      given as M in the specification; causes the
FT      amino acid substitution L91M"
FT      variation
FT      replace (1984, G)
FT      /*tag= x
FT      /label= PS14
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as R in the specification; causes the
FT      amino acid substitution H94R"
FT      variation
FT      replace (1997, G)
FT      /*tag= s
FT      /label= PS15
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as S in the specification"
FT      variation
FT      replace (2014, C)
FT      /*tag= t
FT      /label= PS16
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as Y in the specification; causes the
FT      amino acid substitution V104A"
FT      variation
FT      replace (2022, T)
FT      /*tag= u
FT      /label= PS17
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as W in the specification; together
FT      with PS18 causes the amino acid substitution
FT      T107F"
FT      variation
FT      replace (2023, T)
FT      /*tag= v
FT      /label= PS18
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as Y in the specification; together
FT      with PS17 causes the amino acid substitution
FT      T107F"
FT      variation
FT      replace (2028, G)
FT      /*tag= w
FT      /label= PS19
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as R in the specification; causes the
FT      amino acid substitution I109Y"
FT      variation
FT      replace (2036, C)
FT      /*tag= x

```

```

FT      variation      /label= PS20
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as Y in the specification"
FT      replace (2039, T)
FT      /*tag= y
FT      /label= PS21
FT      /note= "Known single nucleotide polymorphism (SNP);
FT      given as Y in the specification"
FT      intron
FT      2056..2605
FT      /*tag= z
FT      /number= 2
FT      /cons_splice= (5'site:NO, 3'site:YES)
FT      replace (2062, G)
FT      /*tag= aa
FT      /label= PS22
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as R in the specification"
FT      variation
FT      replace (2067, G)
FT      /*tag= ab
FT      /label= PS23
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as K in the specification"
FT      variation
FT      replace (2118, T)
FT      /*tag= ac
FT      /label= PS24
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as Y in the specification"
FT      variation
FT      replace (2170, A)
FT      /*tag= ad
FT      /label= PS25
FT      /note= "Known single nucleotide polymorphism (SNP);
FT      given as R in the specification"
FT      variation
FT      replace (2179, C)
FT      /*tag= ae
FT      /label= PS26
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as S in the specification"
FT      exon
FT      2606..2758
FT      /*tag= af
FT      /number= 3
FT      /label= PS27
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as W in the specification; causes the
FT      amino acid substitution F120I"
FT      variation
FT      replace (2635, C)
FT      /*tag= ah
FT      /label= PS28
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as Y in the specification; causes the
FT      amino acid substitution W128R"
FT      variation
FT      replace (2659, A)
FT      /*tag= ai
FT      /label= PS29
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as R in the specification; together
FT      with PS30 causes the amino acid substitution
FT      V136I"
FT      variation      replace (2661, C)

```

```

Query Match      39.3%; Score 660; DB 24; Length 6472;
Best Local Similarity 99.3%; Pred. No. 1.4e-237;
Matches 1010; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 621 ACCCTCAGGCACCGCGGAGGATGTTGTCACAGGCTGGGCAAGGGCTTCGGGTACCA 680
    |
Db 2 ACCCTCAGGCACCGCGGAGGATGTTGTCACAGGCTGGGCAAGGGCTTCGGGTACCA 61
    |
Qy 681 ACTGGAGCTCTGGCAACAGCCCTGTTGCAAAACAAGAGCCATAGCCGCCGACAGCCCA 740
    |
Db 62 ACTGGAGCTCTGGGACAGCCCTGTTGCAAAACAAGAGCCATAGCCGCCGACAGCCCA 121

```







CC The nucleotide sequence of the cDNA coding region for the human  
CC auxillary cytochrome P450 species 2D6 variant 3. The gene contains a  
CC change at base 1457 from C to G as compared to the wild type sequence  
CC (AA087729). This changes the amino acid residue 296 from Thr to Ser. The  
CC cDNA was amplified by PCR using the primers AA087763-6. The product was  
CC cloned into the yeast expression vectors pAAH5N or pAHR to produce the  
CC vectors p2D6 variant 3 for the expression of the cytochrome P450 alone  
CC or p2D6R variant 3 for co-expression with the yeast NADPH-P450  
CC reductase. The vectors are used in a method for evaluating the safety of  
CC a chemical compound by reacting the chemical compound with recombinantly  
CC produced human cytochrome P450 molecular species 1A2 (AA087714), 2C9  
CC (AA087715), 2E1 (AA087716), or 3A4 (AA087717) or their auxillary species  
CC and variants (AA087718-32), and yeast NADPH-P450 reductase, either as a  
CC fused protein or in cell extracts, and analysing the resulting metabolite  
CC to assess the safety of the chemical compound. The method is useful for  
CC determining whether the chemical compound, or its metabolite, will be  
CC converted into a carcinogenic or mutagenic form through metabolism in the  
CC liver.

XX Sequence 1494 BP; 248 A; 507 C; 447 G; 292 T; 0 other;

Query Match 3.6%; Score 61; DB 16; Length 1494;

Best Local Similarity 100.0%; Pred. No. 1.5e-13;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACGACTGGTGGCCCGTGATAGTGCCATCTCCTGCTCCTGGTG 1679

|||||  
Db 1 ATGGGGCTAGAACGACTGGTGGCCCGTGATAGTGCCATCTCCTGCTCCTGGTG 60

QY 1680 G 1680

||  
Db 61 G 61

RESULT 12

AAT28395

ID AAT28395 standard; DNA; 1494 BP.

XX AAT28395;

AC AAT28395;

DT 11-OCT-1996 (first entry)

DE Human cytochrome P450 molecular species 2D6 variant #1 gene.

XX Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;  
XX liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;  
KW evaluation; safety; fusion protein; metabolite; detoxification;  
KW carcinogenic; ds.

XX Homo sapiens.

OS JP08056695-A.

PN 05-MAR-1996.

PD 15-JUL-1994;

PF 94JP-0164184.

XX 17-JUN-1994;

PR 94JP-0136053.

PR 20-JUL-1993;

PR 93JP-0201120.

PR 30-JUL-1993;

XX 93JP-0208279.

XX (SUMO ) SUMITOMO CHEM CO LTD.

XX WPI; 1996-182311/19.

DR P-PSDB; AAR93182.

XX Novel method for the evaluation of the safety of a cpd. - using a

PT human cytochrome P450 and yeast NADPH reductase to determine whether

PT the analyte cpd. is detoxified or metabolised to a carcinogen

XX Example 1; Page 49-51; 74pp; Japanese.

PS

XX

CC This is the nucleotide sequence of the human cytochrome P450 molecular  
CC species 2D6 variant #1 which encodes a protein of 497 amino acids. The  
CC gene was amplified from a human liver derived cDNA library as 2  
CC fragments of 0.4 and 0.9 kb using primers AAT26953-6. The prod. was  
CC cloned into the yeast expression vector pAAH5N to generate plasmid p2D6  
CC for prodn. of the cytochrome only or into the vector pAHR to generate  
CC the plasmid p2C6R for co-prodn. with the yeast NADPH-P450 reductase. The  
CC sequence is placed under control of the yeast ADH gene promoter and  
CC terminator. The vectors are used in a method for evaluating the safety of  
CC a cpd. by reacting the test cpd. with recombinantly produced human  
CC cytochrome P450 mol. species 1A2 (AAT28380), 2C9 (AAT28381), 2E1  
CC (AAT28382), 3A4 (AAT28383) or their variants (AAT28384-98) together with  
CC yeast NADPH-P450 reductase (either as a fused protein or as a cell  
CC extract) and analysing the resultant metabolite. The cpd. is considered  
CC "safe" if it is detoxified or not rendered carcinogenic or "unsafe" if it  
CC is not detoxified or is metabolised to a carcinogenic cpd.

SQ Sequence 1494 BP; 248 A; 508 C; 446 G; 292 T; 0 other;

Query Match 3.6%; Score 61; DB 17; Length 1494;

Best Local Similarity 100.0%; Pred. No. 1.5e-13;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACGACTGGTGGCCCGTGATAGTGCCATCTCCTGCTCCTGGTG 1679

|||||  
Db 1 ATGGGGCTAGAACGACTGGTGGCCCGTGATAGTGCCATCTCCTGCTCCTGGTG 60

QY 1680 G 1680

||  
Db 61 G 61

RESULT 13

AAT28396

ID AAT28396 standard; DNA; 1494 BP.

XX AAT28396;

AC AAT28396;

DT 11-OCT-1996 (first entry)

DE Human cytochrome P450 molecular species 2D6 variant #2 gene.

XX Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;  
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;  
KW evaluation; safety; fusion protein; metabolite; detoxification;  
KW carcinogenic; ds.

XX Homo sapiens.

OS JP08056695-A.

PN 05-MAR-1996.

PD 15-JUL-1994;

PF 94JP-0164184.

XX 17-JUN-1994;

PR 94JP-0136053.

PR 20-JUL-1993;

PR 93JP-0201120.

PR 30-JUL-1993;

XX 93JP-0208279.

XX (SUMO ) SUMITOMO CHEM CO LTD.

XX WPI; 1996-182311/19.

DR P-PSDB; AAR93183.

XX Novel method for the evaluation of the safety of a cpd. - using a

PT human cytochrome P450 and yeast NADPH reductase to determine whether

PT the analyte cpd. is detoxified or metabolised to a carcinogen

XX Example 1; Page 51-53; 74pp; Japanese.

PS This is the nucleotide sequence of the human cytochrome P450 molecular

XX species 2D6 variant #2 which encodes a protein of 497 amino acids. The

CC gene was amplified from a human liver derived cDNA library as 2

CC fragments of 0.4 and 0.9 kb using primers AAT26953-6. The prod. was



CC cloned into the yeast expression vector pAAH5N to generate plasmid p2D6  
CC for prodn. of the cytochrome only or into the vector pAHRR to generate  
CC the plasmid p2C6R for co-prodn. with the yeast NADPH-P450 reductase. The  
CC sequence is placed under control of the yeast ADH gene promoter and  
CC terminator. The vectors are used in a method for evaluating the safety of  
CC a cpd. by reacting the test cpd. with recombinantly produced human  
CC cytochrome P450 mol. species 1A2 (AAT28380), 2C9 (AAT28381), 2E1  
CC (AAT28382), 3A4 (AAT28383) or their variants (AAT28384-98) together with  
CC yeast NADPH-P450 reductase (either as a fused protein or as a cell  
CC extract) and analysing the resultant metabolite. The cpd. is considered  
CC "safe" if it is detoxified or not rendered carcinogenic or "unsafe" if it  
CC is not detoxified or is metabolised to a carcinogenic cpd.  
SQ Sequence 1494 BP; 248 A; 509 C; 446 G; 291 T; 0 other;

Query Match 3.6%; Score 61; DB 17; Length 1494;

Best Local Similarity 100.0%; Pred. No. 1.5e-13;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTGATAGTGCCCATCTTCCTGCTCCTGGTG 1679

Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTGATAGTGCCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680

Db 61 G 61

RESULT 14

AAT28397

ID AAT28397 standard; DNA; 1494 BP.

AC AAT28397;

XX

DT 11-OCT-1996 (first entry)

DE Human cytochrome P450 molecular species 2D6 variant #3 gene.

XX

KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;  
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;  
KW evaluation; safety; fusion protein; metabolite; detoxification;  
KW carcinogenic; ds.

XX Homo sapiens.

XX JP08056695-A.

XX 05-MAR-1996.

XX 15-JUL-1994; 94JP-0164184.

XX 17-JUN-1994; 94JP-0136053.

XX 20-JUL-1993; 93JP-0201120.

XX 30-JUL-1993; 93JP-0208279.

XX (SUMO ) SUMITOMO CHEM CO LTD.

XX WPI; 1996-182311/19.

XX P-PSDB; AAR93184.

PT Novel method for the evaluation of the safety of a cpd. - using a

PT human cytochrome P450 and yeast NADPH reductase to determine whether

PT the analyte cpd. is detoxified or metabolised to a carcinogen

XX Example 1; Page 53-55; 74pp; Japanese.

XX This is the nucleotide sequence of the human cytochrome P450 molecular  
CC species 2D6 variant #3 which encodes a protein of 497 amino acids. The  
CC gene was amplified from a human liver derived cDNA library as 2  
CC fragments of 0.4 and 0.9 kb using primers AAT26953-6. The prod. was  
CC cloned into the yeast expression vector pAAH5N to generate plasmid p2D6  
CC for prodn. of the cytochrome only or into the vector pAHRR to generate  
CC the plasmid p2C6R for co-prodn. with the yeast NADPH-P450 reductase. The  
CC sequence is placed under control of the yeast ADH gene promoter and

CC terminator. The vectors are used in a method for evaluating the safety of  
CC a cpd. by reacting the test cpd. with recombinantly produced human  
CC cytochrome P450 mol. species 1A2 (AAT28380), 2C9 (AAT28381), 2E1  
CC (AAT28382), 3A4 (AAT28383) or their variants (AAT28384-98) together with  
CC yeast NADPH-P450 reductase (either as a fused protein or as a cell  
CC extract) and analysing the resultant metabolite. The cpd. is considered  
CC "safe" if it is detoxified or not rendered carcinogenic or "unsafe" if it  
CC is not detoxified or is metabolised to a carcinogenic cpd.  
SQ Sequence 1494 BP; 248 A; 508 C; 447 G; 291 T; 0 other;

Query Match 3.6%; Score 61; DB 17; Length 1494;

Best Local Similarity 100.0%; Pred. No. 1.5e-13;

Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTGATAGTGCCCATCTTCCTGCTCCTGGTG 1679

Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTGATAGTGCCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680

Db 61 G 61

RESULT 15

AAT28398

ID AAT28398 standard; DNA; 1494 BP.

XX AAT28398;

XX 11-OCT-1996 (first entry)

DE Human cytochrome P450 molecular species 2D6 variant #4 gene.

KW Human cytochrome P450; amplified; PCR; polymerase chain reaction; primer;  
KW liver; yeast; expression vector; NADPH-P450 reductase; ADH gene promoter;  
KW evaluation; safety; fusion protein; metabolite; detoxification;  
KW carcinogenic; ds.

XX Homo sapiens.

XX JP08056695-A.

XX 05-MAR-1996.

XX 15-JUL-1994; 94JP-0164184.

XX 17-JUN-1994; 94JP-0136053.

XX 20-JUL-1993; 93JP-0201120.

XX 30-JUL-1993; 93JP-0208279.

XX (SUMO ) SUMITOMO CHEM CO LTD.

XX WPI; 1996-182311/19.

XX P-PSDB; AAR93185.

PT Novel method for the evaluation of the safety of a cpd. - using a

PT human cytochrome P450 and yeast NADPH reductase to determine whether

PT the analyte cpd. is detoxified or metabolised to a carcinogen

XX Example 1; Page 55-57; 74pp; Japanese.

XX This is the nucleotide sequence of the human cytochrome P450 molecular  
CC species 2D6 variant #4 which encodes a protein of 497 amino acids. The  
CC gene was amplified from a human liver derived cDNA library as 2  
CC fragments of 0.4 and 0.9 kb using primers AAT26953-6. The prod. was  
CC cloned into the yeast expression vector pAAH5N to generate plasmid p2D6  
CC for prodn. of the cytochrome only or into the vector pAHRR to generate  
CC the plasmid p2C6R for co-prodn. with the yeast NADPH-P450 reductase. The  
CC sequence is placed under control of the yeast ADH gene promoter and  
CC terminator. The vectors are used in a method for evaluating the safety of  
CC a cpd. by reacting the test cpd. with recombinantly produced human  
CC cytochrome P450 mol. species 1A2 (AAT28380), 2C9 (AAT28381), 2E1  
CC (AAT28382), 3A4 (AAT28383) or their variants (AAT28384-98) together with

CC yeast NADPH-P450 reductase (either as a fused protein or as a cell  
 CC extract) and analysing the resultant metabolite. The cpd. is considered  
 CC "safe" if it is detoxified or not rendered carcinogenic or "unsafe" if it  
 CC is not detoxified or is metabolised to a carcinogenic cpd.  
 SQ Sequence 1494 BP; 248 A; 507 G; 447 G; 292 T; 0 other;

Query Match 3.6%; Score 61; DB 17; Length 1494;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-13;  
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACGACACTGGTCCCTGGCGCTGATAGTGGCCATCTTCCTGCTCCTGGTG 1679  
 Db 1 ATGGGGCTAGAACGACACTGGTCCCTGGCGCTGATAGTGGCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680  
 Db 61 G 61

RESULT 16  
 AAT17388  
 ID AAT17388 standard; cDNA; 1494 BP.  
 AC AAT17388;  
 XX 01-AUG-1996 (first entry)  
 DT Human derived cytochrome P4502D6 cDNA.  
 DE Human derived cytochrome; P4502D6; commercial cDNA library; yeast;  
 KW transfection; recombinant production; expression vector; mammal;  
 KW immunisation; sensitisation; antibody; determination; detection;  
 KW non-cross reactive; ds.  
 XX Homo sapiens.

Key Location/Qualifiers  
 CDS 1..1494  
 /\*tag= a  
 JF08027199-A.  
 30-JAN-1996.  
 15-JUL-1994; 94JP-0164186.  
 15-JUL-1994; 94JP-0164186.  
 (SUMO ) SUMITOMO CHEM CO LTD.  
 WPI; 1996-136340/14.  
 P-PSDB; AAR81462.

Antibody recognising human derived cytochrome P4502D6 - allows  
 specific detection of cytochrome P450 species in humans  
 Example 1; Pages 11-13; 13pp; Japanese.

The present sequence encodes the human derived cytochrome (HDC)  
 P4502D6, which was obt'd. from a commercial cDNA library. Yeast  
 were transfected with an expression vector contg. the HDC cDNA,  
 cultured and then disrupted to give a microsomal fraction. The  
 HDC was purified from the fraction, and used to immunise and  
 sensitise a mammal. Blood was drawn from the mammal, and an  
 anti-HDC antibody isolated. The antibody obtd. recognises HDC  
 P4502D6, partic. at a serum dilution rate of 1:10000, and is  
 substantially without cross reaction to other HDC P450 spp..

Query Match 3.6%; Score 61; DB 17; Length 1494;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-13;  
 Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACGACACTGGTCCCTGGCGCTGATAGTGGCCATCTTCCTGCTCCTGGTG 1679  
 Db 1 ATGGGGCTAGAACGACACTGGTCCCTGGCGCTGATAGTGGCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680  
 Db 61 G 61

RESULT 17  
 ABQ72216  
 ID ABQ72216 standard; cDNA; 1494 BP.  
 XX ABQ72216;  
 XX 02-SEP-2002 (first entry)  
 DT Human CYP2D6 gene coding sequence, SEQ ID NO:2.  
 DE Human CYP2D6  
 XX Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;  
 KW chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;  
 KW antiarrhythmic; arrhythmia; adrenoreceptor antagonist; hyperlensin;  
 KW tricyclic antidepressant; procainamide; drug induced lupus syndrome;  
 KW environmentally linked disease; Parkinson's disease; haplotyping;  
 KW genotyping; haplotype; genetic variant; single nucleotide polymorphism;  
 KW SNP; drug screening; drug discovery; gene; ss.

Homo sapiens.  
 XX Key Location/Qualifiers  
 FH 1..1494  
 CDS /\*tag= a  
 /\*product= "CYP2D6"  
 /\*replace (19, A)  
 /\*tag= b  
 /\*label= PS7  
 /\*note= "Known single nucleotide polymorphism (SNP);  
 causes the amino acid substitution V7M"  
 FT variation  
 FT replace (31, A)  
 FT /\*tag= c  
 FT /\*label= PS8  
 FT /\*note= "Known single nucleotide polymorphism (SNP);  
 causes the amino acid substitution V11M"  
 FT variation  
 FT replace (100, T)  
 FT /\*tag= d  
 FT /\*label= PS9  
 FT /\*note= "Known single nucleotide polymorphism (SNP);  
 causes the amino acid substitution P34S"  
 FT variation  
 FT replace (263, A)  
 FT /\*tag= e  
 FT /\*label= PS12  
 FT /\*note= "Novel single nucleotide polymorphism (SNP);  
 causes the amino acid substitution R88H"  
 FT variation  
 FT replace (271, A)  
 FT /\*tag= f  
 FT /\*label= PS13  
 FT /\*note= "Known single nucleotide polymorphism (SNP);  
 causes the amino acid substitution L91M"  
 FT variation  
 FT replace (281, G)  
 FT /\*tag= g  
 FT /\*label= PS14  
 FT /\*note= "Novel single nucleotide polymorphism (SNP);  
 causes the amino acid substitution H94R"  
 FT variation  
 FT replace (294, G)  
 FT /\*tag= h  
 FT /\*label= PS15  
 FT /\*note= "Novel single nucleotide polymorphism (SNP);  
 causes the amino acid substitution V104A"  
 FT variation  
 FT replace (311, C)  
 FT /\*tag= i  
 FT /\*label= PS16  
 FT /\*note= "Novel single nucleotide polymorphism (SNP);  
 causes the amino acid substitution V104A"

```

FT variation replace (319, T)
FT /*tag= j
FT /label= PS17
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS18 causes the amino acid
FT substitution I107F"
FT
FT variation replace (320, T)
FT /*tag= k
FT /label= PS18
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS17 causes the amino acid
FT substitution I107F"
FT
FT variation replace (325, G)
FT /*tag= l
FT /label= PS19
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution I109V"
FT
FT variation replace (333, C)
FT /*tag= m
FT /label= PS20
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT
FT variation replace (336, T)
FT /*tag= n
FT /label= PS21
FT /note= "Known single nucleotide polymorphism (SNP)"
FT
FT variation replace (358, A)
FT /*tag= o
FT /label= PS27
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution F120I"
FT
FT variation replace (382, C)
FT /*tag= p
FT /label= PS28
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution W128R"
FT
FT variation replace (406, A)
FT /*tag= q
FT /label= PS29
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS30 causes the amino acid
FT substitution V136I"
FT
FT variation replace (408, C)
FT /*tag= r
FT /label= PS30
FT /note= "Known single nucleotide polymorphism (SNP);
FT together with PS29 causes the amino acid
FT substitution V136I"
FT
FT variation replace (451, G)
FT /*tag= s
FT /label= PS31
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution Q151E"
FT
FT variation replace (463, A)
FT /*tag= t
FT /label= PS32
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution E155K"
FT
FT variation replace (696, C)
FT /*tag= u
FT /label= PS35
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT
FT variation replace (1012, A)
FT /*tag= v
FT /label= PS36
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V338W"
FT
FT variation replace (1030, T)
FT /*tag= w
FT /label= PS37
FT /note= "Novel single nucleotide polymorphism (SNP);
FT generates a premature stop codon (R344STOP)"
FT
FT variation replace (1083, C)
FT /*tag= x

```

---

```

FT FT /label= PS38
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (1457, C)
FT /*tag= y
FT /label= PS41
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution S486T"
FT
FT variation replace (1489, T)
FT /*tag= z
FT /label= PS42
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution R497C"
FT
XX WO200238589-A2.
XX
XX 16-MAY-2002.
XX
XX 09-NOV-2001; 2001WO-US47396.
XX
XX 09-NOV-2000; 2000US-247943P.
XX
XX (GENA-) GENAISSANCE PHARM INC.
XX
XX Anastasio AE, Chew A, Choi JY, Denton RR, Nandabalan K;
XX Petersen N, Rounds E;
XX WPI; 2002-519292/55.
XX P-PSDB; ABB09563.
XX
XX Novel genetic variants of Cytochrome P450, Subfamily IID, Polypeptide 6
XX isogenes, useful for improving efficiency and reliability in drug
XX development for treating hypertension, arrhythmias and Parkinson's
XX disease.
XX
XX Claim 25; Fig 2; 158pp; English.
XX
XX The invention relates to a method for haplotyping the cytochrome P450,
XX subfamily IID, polypeptide 6 (CYP2D6) gene (ABQ72215, ABQ72364) of an
XX individual, and also describes 29 novel polymorphic sites within the
XX human CYP2D6 gene. The CYP2D6 gene is located on chromosome 22q13.1 and
XX contains 9 exons which encode a 497 amino acid protein (AB09563). CYP2D6
XX is a mono-oxygenase involved in the detoxification of many drugs and
XX environmental chemicals. It plays a role in the metabolism of drugs such
XX as antiarrhythmics, adrenoceptor antagonists and tricyclic
XX antidepressants, and is also involved in the formation of a metabolite
XX linked to the drug-induced lupus syndrome observed with procainamide.
XX Variations in CYP2D6 activity or expression may also influence an
XX individual's susceptibility to environmentally-linked diseases, and it
XX has been demonstrated that CYP2D6 activity may be involved in the
XX pathogenesis of Parkinson's disease, with individuals with a less active
XX form of the enzyme tending to have an earlier onset of this condition.
XX CYP2D6 nucleic acid sequences are useful in studying the expression and
XX function of CYP2D6, and in expressing CYP2D6 protein for use in screening
XX drugs for the treatment of CYP2D6-associated diseases (e.g.,
XX hypertension, atrial and ventricular arrhythmias, Parkinson's disease,
XX and drug-induced lupus syndrome) or which are metabolised by CYP2D6.
XX CYP2D6 nucleic acids and proteins are also useful in studying the effect
XX of polymorphisms on the biological activity of CYP2D6. Polymorphisms
XX in the target region may be determined by the use of allele-specific
XX oligonucleotides (ASOS; ABQ72217-ABQ72303) as probes and primers,
XX and by primer extension using oligonucleotide primers comprising
XX sequences ABQ72304-ABQ72361. The method of the invention is useful for
XX haplotyping the CYP2D6 gene in populations and in individuals, enabling
XX decisions to be made as to whether CYP2D6 is a likely therapeutic target
XX for a disease of interest, and to control for genetically-based bias in
XX the design of drugs that target or are metabolised by CYP2D6. In

```

Query Match 3.6%; Score 61; DB 24; Length 1494;  
Best Local Similarity 100.0%; Pred. No. 1.5e-13;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGATAGTGGCCCTGATAGTGGCCACTTCCTCTCCTGGT 1679  
|||||

Dbl 1 ATGGGGCTAGAACGACTGGTGGCCCTGGCCGTGATAGTGGCCATCTCTCTGCTCGTG 60

Qy 1680 G 1680  
|  
Db 61 G 61

RESULT 18

AAQ12893  
ID AAQ12893 standard; cDNA; 1586 BP.

XX  
AC AAQ12893;

XX 22-OCT-1991 (first entry)

DE Debrisoquine hydroxylase variant "a".

KW db1; Extensive metaboliser/poor metaboliser; EM/PM; ss.

OS Synthetic.

FH Key Location/Qualifiers

FT exon 1..180

FT /\*tag= a

FT /number= 1

FT 181..245

FT /\*tag= b

FT /number= 1

FT /note= "first 64 bases of intron 1"

FT 383

FT /\*tag= c

FT /note= "c.f. C(383) in Gonzalez et al., (1988)

FT Nature, 331, 442-445."

FT 506

FT /\*tag= d

FT /note= "G(506) is deleted; results in frameshift

FT c.f. genomic P450II2D sequence"

XX W09110745-A.

XX 25-JUL-1991.

XX 17-JAN-1991; 91WO-GB000066.

XX 18-JAN-1990; 90GB-0001181.

XX (IMCR ) IMPERIAL CANCER RES.

XX Wolf CR, Miles JS, Spurr NK, Gough AC;

XX WPI; 1991-238033/32.

XX Identification of cytochrome P450 enzyme bufuralol-1'-hydroxylase  
- as indication of extensive-poor metaboliser phenotypes,  
important for drug dosage

XX Example 1; Fig 2; 43pp; English.

XX This full-length db1 variant sequence was obtained from two  
overlapping clones (pMP32 and pMP33). It is not expected to encode  
a functional protein as it contains part of intron 1 and has a  
frameshift relative to the functional gene. The other base pair  
differences between the pMP32/pMP33 sequence and the normal db1  
sequence form the basis of detection methods designed to  
distinguish EM/PM phenotypes. For example, at position 294 of the  
variant sequence G replaces the wild-type C. This introduces a  
HaeIII restriction site in PM individuals. PCR amplification of a  
172bp fragment containing this site, followed by HaeIII digestion  
allows homozygous EM individuals (who lack this site) to be  
distinguished from PM individuals. See AAQ12885-Q12894.

XX Sequence 1586 BP; 286 A; 527 C; 464 G; 309 T; 0 other;

Query Match 3.6%; Score 61; DB 12; Length 1586;  
Best Local Similarity 100.0%; Pred. No. 1.5e-13;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1620 ATGGGGCTAGAACGACTGGTGGCCCTGGCCGTGATAGTGGCCATCTCTCTGCTCGTG 1679  
|  
Db 1 ATGGGGCTAGAACGACTGGTGGCCCTGGCCGTGATAGTGGCCATCTCTCTGCTCGTG 60

Qy 1680 G 1680  
|  
Db 61 G 61

RESULT 19

ABAL6633/c

ID ABAL6633 standard; DNA; 13836 BP.

XX  
AC ABAL6633;

XX 23-JAN-2002 (first entry)

DE Human nervous system related polynucleotide SEQ ID NO 8964.

XX

KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

XX Homo sapiens.

XX OS

XX PN W0200159063-A2.

XX 16-AUG-2001.

XX 17-JAN-2001; 2001WO-US01334.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

XX 14-AUG-2000; 2000US-0225268.

XX 14-AUG-2000; 2000US-0225270.

XX 14-AUG-2000; 2000US-0225447.

XX 14-AUG-2000; 2000US-0225757.

XX 14-AUG-2000; 2000US-0225758.

XX 14-AUG-2000; 2000US-0225759.

XX 18-AUG-2000; 2000US-0226279.

XX 22-AUG-2000; 2000US-0226681.

XX 22-AUG-2000; 2000US-0226886.

XX 22-AUG-2000; 2000US-0227182.



ABK93054  
ID ABK93054 standard; cDNA; 303 BP.  
XX  
AC ABK93054;  
XX  
DT 22-AUG-2002 (first entry)  
XX  
DE cDNA encoding human cytochrome p450 db1.  
XX  
KW Human; cancer; DNA array; gene expression analysis; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN JP2002058495-A.  
XX  
PD 26-FEB-2002.  
XX  
PF 22-AUG-2000; 2000JP-0255737.  
XX  
PR 22-AUG-2000; 2000JP-0255737.  
XX  
PA (HITA ) HITACHI LTD.  
XX  
DR WPI; 2002-397522/43.  
XX  
PT A DNA array that observes the expression of cancer related genes is  
PT useful for the elucidation of the action mechanism of anticancer agents  
PT  
XX  
PS Disclosure; Page 40; 71pp; Japanese.  
XX  
CC The invention relates to a DNA array for identifying cancer related  
CC genes. The DNA array comprises: (a) statistically significantly higher  
CC results of over 10% in homology search of a DNA molecule describing  
CC different genes or a complementary DNA (cDNA); (b) statistically  
CC significantly higher results (p value) of over 10% in homology search  
CC with human Alu sequence; (c) difference of melting temperature of the  
CC DNA fragment and hybridisation temperature at 30 degree C or less; (d)  
CC the DNA fragment prepared by immobilisation of a DNA short sequence with  
CC a specific higher order structure; and (e) the DNA fragment without  
CC short repetitive sequence. The array is used in elucidation of the action  
CC mechanism of anticancer agents. ABK92989-ABK93090 represent human  
CC cancer related genes and associated primers of the invention.  
XX  
SQ Sequence 303 BP; 42 A; 109 C; 100 G; 52 T; 0 other;  
  
Query Match 3.1%; Score 52; DB 24; Length 303;  
Best Local Similarity 100.0%; Pred. No. 4.5e-10;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1629 GAAGCACTGGTGGCCCGCGGTGATAGTGGCCATCTTCCTGCTCGGTGG 1680  
Db 1 GAAGCACTGGTGGCCCGCGGTGATAGTGGCCATCTTCCTGCTCGGTGG 52  
  
RESULT 21  
AAK77606/c  
ID AAK77606 standard; DNA; 6801 BP.  
XX  
AC AAK77606;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32418.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 28-JUN-2000; 2000US-0209467.  
PR 30-JUN-2000; 2000US-0214886.  
PR 07-JUL-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0217496.  
PR 26-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0227009.  
PR 01-SEP-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.



PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225737.  
 PR 14-AUG-2000; 2000US-0225738.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 08-SEP-2000; 2000US-02311242.  
 PR 08-SEP-2000; 2000US-02311243.  
 PR 08-SEP-2000; 2000US-02311244.  
 PR 08-SEP-2000; 2000US-0231143.  
 PR 08-SEP-2000; 2000US-0231144.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 12-SEP-2000; 2000US-0232081.  
 PR 14-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 21-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234423.  
 PR 25-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 13-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 20-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 08-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 05-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 32419; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patients own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I) by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703



CC to AAK87694 represent human immune/haematopoietic antigen genomic.  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 6801 BP; 1630 A; 1880 C; 1686 G; 1605 T; 0 other;

Query Match 3.0%; Score 50; DB 22; Length 6801;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 218 CACCTGTAATCCAGCTACTTAGGAGCTGAGCAGGAGAGATTGCTTGAA 267  
|||||  
DB 138 CACCTGTAATCCAGCTACTTAGGAGCTGAGCAGGAGAGATTGCTTGAA 89

RESULT 23  
AAL04340/C  
ID AAL04340 standard; DNA; 32194 BP.

XX AAL04340;

DT 21-NOV-2001 (first entry)

XX Human reproductive system related antigen DNA SEQ ID NO: 7028.

XX Human reproductive system related antigen; reproductive system disorder;  
KW cancer; gene therapy; ds.

XX Homo sapiens.

XX WO200155320-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01339.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 11-JUL-2000; 2000US-0217487.

XX 14-JUL-2000; 2000US-0217496.

XX 26-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

XX 14-AUG-2000; 2000US-0225213.

XX 14-AUG-2000; 2000US-0225214.

XX 14-AUG-2000; 2000US-0225266.

XX 14-AUG-2000; 2000US-0225267.

PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI; 2001-465570/50.  
XX  
XX Isolated nucleic acid molecule encoding a reproductive system antigen  
PT is used in preventing, treating or ameliorating a medical condition -  
PT  
XX Disclosure; SEQ ID NO 7028; 1297pp + Sequence Listing; English.  
PS  
XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention.  
XX  
SQ Sequence 32194 BP; 7448 A; 8292 C; 8454 G; 8000 T; 0 other;  
Query Match 3.0%; Score 50; DB 22; Length 32194;  
Best Local Similarity 100.0%; Pred.No.1.le-09;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 218 CACCTGTAATCCCGAGTACTTAGGAGGCTGAGCGAGAGAAATTCCTTGA 267  
|||||  
Db 10677 CACCTGTAATCCCGAGTACTTAGGAGGCTGAGCGAGAGAAATTCCTTGA 10628  
RESULT 24  
AAL36886/c  
ID AAL36886 standard; DNA; 32177 BP.  
XX  
AC AAL36886;  
XX  
XX 08-JAN-2002 (first entry)  
XX  
XX Human musculoskeletal system related polynucleotide SEQ ID NO 3251.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system; ds.  
XX  
XX Homo sapiens.  
OS

XX  
PN WO200155367-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01338.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.



PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225113.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226688.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235634.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI; 2001-465570/50.  
 XX  
 PT Isolated nucleic acid molecule encoding a reproductive system antigen -  
 PT is used in preventing, treating or ameliorating a medical condition -  
 XX  
 PS Disclosure; SEQ ID NO 6967; 1297pp + Sequence Listing; English.  
 XX  
 CC The present invention provides the protein and coding sequences of a  
 CC number of human reproductive system related antigens. These can be used  
 CC in the prevention and treatment of reproductive system disorders,  
 CC including cancer. The present sequence is a genomic sequence encoding a  
 CC protein of the invention.  
 XX  
 SQ Sequence 32177 BP; 9305 A; 5927 C; 6437 G; 10508 T; 0 other;



PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
XX Disclosure; SEQ ID NO 36984; 3071pp + Sequence Listing; English.  
PS  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 38136 BP; 11151 A; 6962 C; 7661 G; 12362 T; 0 other;  
Query Match 2.9%; Score 49; DB 22; Length 38136;  
Best Local Similarity 100.0%; Pred. No. 2.6e-09;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 138 CAGGAGTTCAGACTAGCCTGGCCACATGGTGAACCCCTATCTCTACT 186  
|||||||  
Db 3375 CAGGAGTTCAGACTAGCCTGGCCACATGGTGAACCCCTATCTCTACT 3327  
RESULT 27  
AAK82171/c  
ID AAK82171 standard; DNA; 38140 BP.  
XX  
XX AAK82171;  
AC  
XX  
XX  
DT 07-NOV-2001 (first entry)  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36983.  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW

KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX Homo sapiens.  
OS  
XX WO200157182-A2.  
PN  
XX 09-AUG-2001.  
PD  
XX 17-JAN-2001; 2001WO-US01354.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0217496.  
PR 26-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.



PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0228928.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2000US-0254097.  
PR 05-JAN-2001; 2000US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases -

Disclosure: SEQ ID NO 13036; 170lpp + Sequence Listing; English.

The invention relates to novel genes (ABAI1004-ABA21534) and proteins (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating



CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 101 BP; 23 A; 20 C; 38 G; 20 T; 0 other;

Query Match 2.9%; Score 48; DB 22; Length 101;  
Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGGAATTGCTTGAA 267  
|||||  
Db 28 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGGAATTGCTTGAA 75

RESULT 29

ABA20707  
ID ABA20707 standard; DNA; 101 BP.

XX  
AC ABA20707;

XX  
DT 23-JAN-2002 (first entry)

XX  
DE Human nervous system related polynucleotide SEQ ID NO 13038.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
KW antiparkinsonian; antiskickling; antianaemic; antiarthritic; cancer;  
KW antihumetic; hepatotropic; cerebroprotective; antinflammatory;  
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

OS Homo sapiens.

PN WO200159063-A2.

XX  
PD 16-AUG-2001.

XX  
PF 17-JAN-2001; 2001WO-US01334.

XX  
PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0242221.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.



PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 23-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.

PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251858.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
PI  
XX  
DR WPI; 2001-483426/52.  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
PS Disclosure; SEQ ID NO 20153; 3071pp + Sequence Listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 1627 BP; 481 A; 286 C; 333 G; 527 T; 0 other;  
Query Match 2.9%; Score 48; DB 22; Length 1627;  
Best Local Similarity 100.0%; Pred. No. 1.1e-08;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 220 CCTGTAATCCAGCTACTTAGGAGCTGAGGACGAGAGATTGCTTGA 267  
|||||  
Db 1197 CCTGTAATCCAGCTACTTAGGAGCTGAGGACGAGAGATTGCTTGA 1244  
RESULT 31  
AAK77234/C  
ID AAK77234 standard; DNA; 1627 BP.  
XX  
AC AAK77234;  
XX

DT 07-NOV-2001 (first entry)  
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32046.  
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
KW  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.



XX 08-FEB-2001 (first entry)

XX Polymorphic repeat microsatellite sequences present in the CTIA4 locus.

DE Autoimmune disease; polymorphic microsatellite repeat; PMR; CD28 gene;

XX ICOS gene; CTIA4 gene; costimulatory receptor gene locus; CGRL; lupus;

KW insulin-dependent diabetes mellitus; IDDM; Addison's disease; leprosy;

KW Graves disease; autoimmune hypothyroidism; myasthenia gravis; thymoma;

KW thyroiditis; postpartum thyroiditis; rheumatoid arthritis;

KW Hashimoto's disease; coeliac disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX satellite 5722...5746

FT /\*tag= a

FT /note= "sara41/42 microsatellite repeat"

FT 6550...6597

FT /\*tag= b

FT /note= "sara43/44 microsatellite repeat"

FT 1911...1956

FT /\*tag= c

FT /note= "PW210/211 microsatellite repeat"

FT 23904...23957

FT /\*tag= d

FT /note= "sara45/46 microsatellite repeat"

FT 27689...27780

FT /\*tag= e

FT /note= "sara17/18 microsatellite repeat"

FT 30766...30801

FT /\*tag= f

FT /note= "sara19/20 microsatellite repeat"

XX WO200056856-A2.

XX 28-SEP-2000.

XX 24-MAR-2000; 2000WO-US07938.

XX 25-MAR-1999; 99US-0126215.

XX (GEMY ) GENETICS INST INC.

XX Ling V, Wu P, Gray GS;

XX WPI; 2000-628257/60.

XX Determining predisposition of humans to develop autoimmune disease

XX Involves detecting polymorphic microsatellite repeat sequence within

XX human costimulatory receptor gene locus -

XX Claim 2; Page 67-82; 160pp; English.

XX Two human bacterial artificial chromosome (BAC) clones that included

XX and flanked the human CTIA4 locus were cloned and sequenced. The

XX sequence data was assembled into a contiguous sequence that is presented

XX in AAA96363-68. AAA96363-64 comprise BAC clone 22700, and AAA96365-68

XX comprise BAC clone 22608. The sequences contain polymorphic

XX microsatellite repeat (PMR) sequences. The specification describes a

XX method for determining the predisposition of a human subject to develop

XX autoimmune disease. The method comprises detecting a PMR sequence in the

XX CD28, ICOS gene or CTIA4 gene of the human costimulatory receptor gene

XX locus (hCGRL). PMR sequences vary in length among individuals and can be

XX amplified to generate products that differ in size. These products can

XX then be detected by rapid and convenient high resolution processes. The

XX method is useful for determining the predisposition of insulin-dependent

XX diabetes mellitus (IDDM), Addison's disease, Graves disease, autoimmune

XX hypothyroidism, myasthenia gravis, thymoma, lupus, thyroiditis,

XX postpartum thyroiditis, rheumatoid arthritis, Hashimoto's disease,

XX coeliac disease and leprosy. PMR sequences within hCGRL are useful as

XX markers in a variety of assays and in the field of forensic medicine,

XX disease diagnosis and human genome mapping.

XX SQ Sequence 50000 BP; 14521 A; 9882 C; 11191 G; 14406 T; 0 other;

XX Query Match 2.7%; Score 45; DB 21; Length 50000;

XX Best Local Similarity 100.0%; Pred. No. 7.8e-08;

XX Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 TGAAGTCAGGAGTTCACAGACTAGCCTGGCCCAACATGTGAAACCC 176

DB 14836 TGAAGTCAGGAGTTCACAGACTAGCCTGGCCCAACATGTGAAACCC 14792

XX RESULT 34

XX AAC89560/C

XX ID AAC89560 standard; DNA; 122186 BP.

XX AC AAC89560;

XX DT 08-MAR-2001 (first entry)

XX Human histone deacetylase HDAC-D coding sequence.

XX Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;

XX HDAC-D; cell cycle; tumorigenesis; cancer; inhibitor; antisense;

XX gene therapy; ds.

XX Homo sapiens.

XX WO200071703-A2.

XX PN 30-NOV-2000.

XX PD 03-MAY-2000; 2000WO-IB01252.

XX PF 03-MAY-1999; 99US-0132287.

XX PR (METH-) METHYLGENE INC.

XX PA Macleod AR, Li Z, Besterman JM;

XX PI WPI; 2001-016407/02.

XX Antisense oligonucleotide that inhibits expression of a histone

XX deacetylase, useful for treating and/or alleviating the symptoms of

XX neoplasia, or for inhibiting neoplastic cell growth in an animal -

XX Disclosure; Page 89-125; 125pp; English.

XX The present invention provides inhibitors of histone deacetylase enzymes

XX such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These

XX inhibitors may be antisense strands or they may be compounds identified

XX by contacting the enzyme with the compound and measuring the resulting

XX enzyme activity. These inhibitors are useful for treating cancers and for

XX identifying which histone deacetylase is involved in a neoplasia.

XX SQ Sequence 122186 BP; 29016 A; 31077 C; 32425 G; 29668 T; 0 other;

XX Query Match 2.6%; Score 44; DB 22; Length 122186;

XX Best Local Similarity 100.0%; Pred. No. 1.6e-07;

XX Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 CACACCTGTATCCAGCTACTTAGGAGCTGAGCGAGAGAGT 259

DB 35936 CACACCTGTATCCAGCTACTTAGGAGCTGAGCGAGAGAGT 35893

XX RESULT 35

XX ABRK3459/C

XX ID ABRK3459 standard; cDNA; 128978 BP.

XX AC ABRK3459;

XX DT 14-AUG-2002 (first entry)



XX AC ABAL7344;  
 XX XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Human nervous system related polynucleotide SEQ ID NO 9675.  
 XX  
 KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
 KW antiparkinsonian; antischlicking; antianaemic; antiarthritic; cancer;  
 KW antineumatic; hepatoprotective; cerebroprotective; antiinflammatory;  
 KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;  
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
 OS Homo sapiens.  
 XX  
 XX WO200159063-A2.  
 XX  
 XX 16-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01334.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 30-JUN-2000; 2000US-0214886.  
 PR 07-JUL-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 11-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 14-JUL-2000; 2000US-0217496.  
 PR 26-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 14-AUG-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0227009.  
 PR 01-SEP-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 05-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 08-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.

PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 21-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 13-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 20-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0242221.  
 PR 08-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0244674.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.





PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX (HYSE-) HYSEQ INC.  
XX  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI; 2001-514838/56.  
DR P-PSDB; AAO13377.  
XX  
XX Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
XX Claim 1; SEQ ID NO 13368; 1399pp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO.  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 387 BP; 93 A; 99 C; 78 G; 117 T; 0 other;  
SQ

Query Match 2.5%; Score 42; DB 22; Length 387;  
Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 TGGGTGGGTGGTGCATGCTATATCCAGCACCTTTGGGAG 111  
Db 105 TGGGTGGGTGGTGCATGCTATATCCAGCACCTTTGGGAG 64  
|||||

RESULT 40  
ABL85143/C  
ID ABL85143 standard; cDNA; 397 BP.  
XX  
XX ABL85143;  
AC  
XX  
XX 17-MAY-2002 (first entry)  
DT  
XX  
XX Human ovarian cancer related cDNA clone SEQ ID NO:8121.  
DE  
XX  
XX Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200192581-A2.  
PN  
XX  
XX 06-DEC-2001.  
PD  
XX  
XX 29-MAY-2001; 2001WO-US17756.  
PF  
XX  
XX 26-MAY-2000; 2000US-207484P.  
PR  
XX  
XX (CORI-) CORIXA CORP.  
PA  
XX  
XX Algate PA, Harlocker SL, Jones R;  
PI  
XX  
XX WPI; 2002-122075/16.  
DR  
XX  
XX Composition for therapy and diagnosis of ovarian cancer comprising  
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
PT polypeptide, antibody specific to polypeptide or T cell expressing

PT polypeptide -  
XX  
XX Claim 1; SEQ ID 8121; 489pp; English.  
XX  
XX The present invention describes a composition (I) comprising: carriers  
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour  
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence  
CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to  
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell  
CC population of (II), or antigen presenting cells that express (II).  
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to  
CC (S1) can be used for detecting ovarian cancer in a patient's biological  
CC sample preferably serum or ovarian tissue. The method comprises  
CC contacting a biological sample from a patient with (IV), detecting the  
CC amount of polynucleotide hybridising to (IV) and comparing the amount to  
CC a predetermined cutoff value and thereby detecting ovarian cancer in the  
CC patient, where the amount of polynucleotide hybridising to (IV) is  
CC detected preferably by polymerase chain reaction (PCR). (I) comprising  
CC (III) and/or (II) is useful for stimulating and/or expanding T cells  
CC specific for an ovarian tumour protein comprising contacting T cells  
CC with (III) or (II). (III) is useful in design and preparation of  
CC ribozyme molecules for inhibiting expression of the tumour polypeptides  
CC and proteins in tumour cells; and to isolate a full length gene from a  
CC suitable library e.g., a tumour cDNA library using well known  
CC techniques.  
XX  
XX Sequence 397 BP; 100 A; 104 C; 84 G; 107 T; 2 other;  
SQ

Query Match 2.5%; Score 42; DB 24; Length 397;  
Best Local Similarity 100.0%; Pred. No. 2.3e-06;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 CACCTGTATATCCAGCTACTTAGGAGGCTGAGGCGAGGAGAT 259  
Db 139 CACCTGTATATCCAGCTACTTAGGAGGCTGAGGCGAGGAGAT 98  
|||||

RESULT 41  
ABV20089  
ID ABV20089 standard; cDNA; 584 BP.  
XX  
XX ABV20089;  
AC  
XX  
XX 13-SEP-2002 (first entry)  
DT  
XX  
XX Human prostate expression marker cDNA 20080.  
DE  
XX  
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200160860-A2.  
PN  
XX  
XX 23-AUG-2001.  
PD  
XX  
XX 20-FEB-2001; 2001WO-US05171.  
PF  
XX  
XX 17-FEB-2000; 2000US-183319P.  
PR  
XX  
XX 16-MAR-2000; 2000US-189862P.  
PR  
XX  
XX 25-MAY-2000; 2000US-207454P.  
PR  
XX  
XX 09-JUN-2000; 2000US-211314P.  
PR  
XX  
XX 18-JUL-2000; 2000US-219007P.  
PR  
XX  
XX 13-DEC-2000; 2000US-255281P.  
PR  
XX  
XX (MTLL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PA  
XX  
XX Schlegel R, Endege WO, Monahan JE;  
PI  
XX  
XX WPI; 2001-662795/76.  
DR  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer -  
PS Claim 1; Page 3276; 11750pp; English.  
XX  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 584 BP; 139 A; 152 C; 129 G; 163 T; 1 other;  
Query Match 2.5%; Score 42; DB 23; Length 584;  
Best Local Similarity 100.0%; Pred. No. 2.2e-06;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 70 TGGGTGCGGTGGCTATGCTATATCCAGCACACTTTGGGAG 111  
DB 242 TGGGTGCGGTGGCTATGCTATATCCAGCACACTTTGGGAG 283  
|||||  
RESULT 42  
ABV25919  
ID ABV25919 standard; cDNA; 584 BP.  
XX  
AC ABV25919;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 25910.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
PA  
XX Schlegel R, Endege WO, Monahan JE;  
PI  
XX WPI; 2001-662795/76.  
XX  
XX Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
XX Claim 1; Page 5240; 11750pp; English.  
PS  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 584 BP; 139 A; 152 C; 129 G; 163 T; 1 other;  
Query Match 2.5%; Score 42; DB 23; Length 584;  
Best Local Similarity 100.0%; Pred. No. 2.2e-06;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 70 TGGGTGCGGTGGCTATGCTATATCCAGCACACTTTGGGAG 111  
DB 242 TGGGTGCGGTGGCTATGCTATATCCAGCACACTTTGGGAG 283  
|||||  
RESULT 43  
AAA81674  
ID AAA81674 standard; DNA; 1180 BP.  
XX  
AC AAA81674;  
XX  
DT 04-DEC-2000 (first entry)  
XX  
DE N. meningitidis partial DNA sequence gnm\_221 SEQ ID NO:221.  
XX  
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;  
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;  
KW Meningococcus B; MenB; ds.  
XX  
OS Neisseria meningitidis.  
XX  
XX WO200022430-A2.  
PN  
XX 20-APR-2000.  
PD  
XX 08-OCT-1999; 99WO-US23573.  
PF  
XX 09-OCT-1998; 98US-0103794.  
PR  
PR 30-APR-1999; 99US-0132068.  
XX  
XX (CHIR ) CHIRON CORP.  
PA  
XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;  
PI Masignani V, Galeotti C, Mora M, Rattl G, Scarscilli M, Scarlato V;  
PI Rappuoli R, Pizza M;  
XX  
XX WPI; 2000-318079/27.  
DR  
XX Isolated nucleotide sequences of Neisseria meningitidis which can be  
PT used in the diagnosis and treatment of N. meningitidis infection and  
PT other Neisserial infections, for example, N.gonorrhoea -  
XX  
XX Claim 7; Page 1535-1536; 1760pp; English.  
PS  
XX The present invention describes methods of obtaining immunogenic  
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414  
CC represent specifically claimed Neisseria meningitidis genomic DNA  
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent  
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to  
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the  
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to  
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF  
CC sequences, which are all used in the exemplification of the present



PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-483426/52.  
XX  
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating cancers and  
XX metastasis -  
XX  
XX  
XX Disclosure: SEQ ID NO 36244; 3071pp + Sequence Listing; English.  
XX  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic  
XX activity, and can be used in gene therapy and vaccine production. (I)  
XX proteins and polynucleotides may be used in the prevention, diagnosis and  
XX treatment of diseases associated with inappropriate (I) expression. For  
XX example, they may be used to treat disorders associated with decreased  
XX expression by rectifying mutations or deletions in a patient's genome  
XX that affect the activity of (I) by expressing inactive proteins or to  
XX supplement the patients own production of (I). Additionally, (I)  
XX polynucleotides may be used to produce the secreted (I), by inserting  
XX the nucleic acids into a host cell and culturing the cell to express the  
XX protein. (I) proteins and polynucleotides may be used to prevent,  
XX diagnose and treat immune/haematopoietic-related diseases, especially  
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
XX to AAK87694 represent human immune/haematopoietic antigen genomic  
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
XX represent sequences used in the exemplification of the present invention.  
XX  
XX Sequence 5690 BP; 1364 A; 1302 C; 1330 G; 1694 T; 0 other;  
SQ

Best Local Similarity 100.0%; Pred. No. 1.5e-06;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 70 TGGTGGGTGGCTCATGCTATATAATCCAGCACACTTTGGGAG 111  
Db 1226 TGGTGGGTGGCTCATGCTATATAATCCAGCACACTTTGGGAG 1267  
RESULT 45  
AAL03263  
ID AAL03263 standard; DNA; 24977 BP.  
XX  
XX AAL03263;  
XX  
XX 21-NOV-2001 (first entry)  
XX  
XX Human reproductive system related antigen DNA SEQ ID NO: 5951.  
XX  
XX Human; reproductive system related antigen; reproductive system disorder;  
XX KW cancer; gene therapy; ds.  
XX  
XX Homo sapiens.  
XX  
XX WO200155320-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01339.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226868.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.



PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225466.  
PR 14-AUG-2000; 2000US-0225467.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 03-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465570/50.

Isolated nucleic acid molecule encoding a reproductive system antigen  
Is used in preventing, treating or ameliorating a medical condition -

Disclosure; SEQ ID NO 8476; 1297pp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a  
number of human reproductive system related antigens. These can be used

CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
XX protein of the invention.

SQ Sequence 24977 BP; 7106 A; 5398 C; 5414 G; 7059 T; 0 other;

Query Match 2.5%; Score 42; DB 22; Length 24977;

Best Local Similarity 100.0%; Pred. No. 1.2e-06;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 CACCTGTAATCCACGCTACTTAGAGGCTGAGGCAGGAGAAAT 259

|||||

Db 24413 CACCTGTAATCCACGCTACTTAGAGGCTGAGGCAGGAGAAAT 24454

RESULT 47

ABL98352

ID ABL98352 standard; DNA; 24977 BP.

XX ABL98352;

DT 21-JUN-2002 (first entry)

DE Human testicular antigen encoding DNA fragment SEQ ID NO: 3004.

XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;  
KW reproductive system disorder; urinary system disorder; gene therapy;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disease; infection; cytostatic; gene; ds.

OS Homo sapiens.

XX WO200155317-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01329.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 18-AUG-2000; 2000US-0225759.

PR 22-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 23-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234984.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0241826.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.









XX	Human testicular antigen encoding DNA fragment SEQ ID NO: 3003.	PR	14-SEP-2000;	2000US-0233063.
DE		PR	14-SEP-2000;	2000US-0233064.
XX		PR	14-SEP-2000;	2000US-0233065.
KW	Human; testicular antigen; testes; cancer; metastasis; immune disorder;	PR	21-SEP-2000;	2000US-0234223.
KW	reproductive system disorder; urinary system disorder; gene therapy;	PR	25-SEP-2000;	2000US-0234274.
KW	cardiovascular disorder; respiratory disorder; neurological disorder;	PR	25-SEP-2000;	2000US-0234997.
KW	gastrointestinal disease; infection; cytostatic; gene; ds.	PR	26-SEP-2000;	2000US-0234998.
XX		PR	26-SEP-2000;	2000US-0235484.
OS		PR	27-SEP-2000;	2000US-0235834.
XX		PR	27-SEP-2000;	2000US-0235836.
PN	WO200155317-A2.	PR	29-SEP-2000;	2000US-0236327.
XX		PR	29-SEP-2000;	2000US-0236367.
XX		PR	29-SEP-2000;	2000US-0236368.
PF		PR	29-SEP-2000;	2000US-0236369.
PF		PR	29-SEP-2000;	2000US-0236370.
XX		PR	02-OCT-2000;	2000US-0236802.
XX		PR	02-OCT-2000;	2000US-0237037.
PR		PR	02-OCT-2000;	2000US-0237038.
PR		PR	02-OCT-2000;	2000US-0237039.
PR		PR	02-OCT-2000;	2000US-0237040.
PR		PR	13-OCT-2000;	2000US-0239935.
PR		PR	13-OCT-2000;	2000US-0239937.
PR		PR	20-OCT-2000;	2000US-0240960.
PR		PR	20-OCT-2000;	2000US-0241221.
PR		PR	20-OCT-2000;	2000US-0241785.
PR		PR	20-OCT-2000;	2000US-0241786.
PR		PR	20-OCT-2000;	2000US-0241787.
PR		PR	20-OCT-2000;	2000US-0241808.
PR		PR	20-OCT-2000;	2000US-0241809.
PR		PR	20-OCT-2000;	2000US-0241826.
PR		PR	01-NOV-2000;	2000US-0246471.
PR		PR	08-NOV-2000;	2000US-0246474.
PR		PR	08-NOV-2000;	2000US-0246475.
PR		PR	08-NOV-2000;	2000US-0246476.
PR		PR	08-NOV-2000;	2000US-0246477.
PR		PR	08-NOV-2000;	2000US-0246478.
PR		PR	08-NOV-2000;	2000US-0246523.
PR		PR	08-NOV-2000;	2000US-0246524.
PR		PR	08-NOV-2000;	2000US-0246525.
PR		PR	08-NOV-2000;	2000US-0246526.
PR		PR	08-NOV-2000;	2000US-0246527.
PR		PR	08-NOV-2000;	2000US-0246528.
PR		PR	08-NOV-2000;	2000US-0246532.
PR		PR	08-NOV-2000;	2000US-0246609.
PR		PR	08-NOV-2000;	2000US-0246610.
PR		PR	08-NOV-2000;	2000US-0246611.
PR		PR	08-NOV-2000;	2000US-0246613.
PR		PR	17-NOV-2000;	2000US-0249207.
PR		PR	17-NOV-2000;	2000US-0249208.
PR		PR	17-NOV-2000;	2000US-0249209.
PR		PR	17-NOV-2000;	2000US-0249210.
PR		PR	17-NOV-2000;	2000US-0249211.
PR		PR	17-NOV-2000;	2000US-0249212.
PR		PR	17-NOV-2000;	2000US-0249213.
PR		PR	17-NOV-2000;	2000US-0249214.
PR		PR	17-NOV-2000;	2000US-0249215.
PR		PR	17-NOV-2000;	2000US-0249216.
PR		PR	17-NOV-2000;	2000US-0249217.
PR		PR	17-NOV-2000;	2000US-0249218.
PR		PR	17-NOV-2000;	2000US-0249244.
PR		PR	17-NOV-2000;	2000US-0249245.
PR		PR	17-NOV-2000;	2000US-0249264.
PR		PR	17-NOV-2000;	2000US-0249265.
PR		PR	17-NOV-2000;	2000US-0249266.
PR		PR	17-NOV-2000;	2000US-0249297.
PR		PR	17-NOV-2000;	2000US-0249299.
PR		PR	17-NOV-2000;	2000US-0249300.
PR		PR	01-DEC-2000;	2000US-0250160.
PR		PR	01-DEC-2000;	2000US-0250391.
PR		PR	05-DEC-2000;	2000US-0251030.
PR		PR	05-DEC-2000;	2000US-0251988.
PR		PR	05-DEC-2000;	2000US-0256719.
PR		PR	06-DEC-2000;	2000US-0251479.
PR		PR	08-DEC-2000;	2000US-0251856.

```
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
DR WPI; 2001-483232/52..  
XX  
XX Nucleic acids encoding 973 human testicular antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating testicular cancer -  
PS Disclosure; SEQ ID NO 3003; 766pp; English.  
XX  
CC The present invention provides the protein and coding sequences of 973  
CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer,  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention.
```

```
XX SQ Sequence 24983 BP; 7104 A; 5397 C; 5409 G; 7073 T; 0 other;
```

```
Query Match          2.5%   Score 42; DB 23; Length 24983;  
Best Local Similarity 100.0%; Pred. No. 1.2e-06;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      218  CACCTGTAAATCCCAGCTACTTTAGGAGGGCTGAGGCAGGAGAAT 259  
           |||||               |              |||||  
Db       24419 CACCTGTAAATCCCAGCTACTTTAGGAGGGCTGAGGCAGGAGAAT 24460
```

```
Search completed: February 11, 2003, 10:33:53  
Job time : 1050 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: February 11, 2003, 06:00:26 ; Search time 2514 Seconds  
(without alignments)  
10822.760 Million cell updates/sec

Title: US-09-942-310-2  
Perfect score: 1680  
Sequence: 1 gaattcaagaccagctgga.....catcttctgctcctgtgg 1680

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 20  
Total number of hits satisfying chosen parameters: 633091

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

EST:*	
1: em_estba.*	BT030898 IL5-MT026
2: em_esthum.*	BF746029 CM4-BT085
3: em_estin.*	BF910677 RC1-UT003
4: em_estmu.*	BF913166 RC1-UT003
5: em_estov.*	B81880 RPC11-1706
6: em_estpl.*	AQ352156 CITBI-EI-
7: em_estro.*	AQ405838 HS_5043_H
8: em_hic.*	AQ309829 CIT-HSP-2
9: gb_est1.*	AQ563791 HS_5338_B
10: gb_est2.*	AA484228 nc89d07.s
11: gb_hic.*	BF677661 602085543
12: gb_est3.*	AQ469676 CITBI-EI-
13: gb_est4.*	BF71528 IL5-IT002
14: gb_est5.*	AQ395268 CITBI-EI-
15: em_estfun.*	AW971963 EST384172
16: em_estom.*	AQ425674 CITBI-EI-
17: gb_gss.*	AQ057132 Pan trogl
18: em_gss_hum.*	AG158320 Pan trogl
19: em_gss_inv.*	AG57369 NBI-875R
20: em_gss_pln.*	B92374 CIT-HSP-217
21: em_gss_vrt.*	BF918155 IL3-UT011
22: em_gss_fun.*	AG099522 Pan trogl
23: em_gss_mam.*	AQ337652 NB6-242C
24: em_gss_mus.*	T47924
25: em_gss_other.*	AA971049 op10h01.s
26: em_gss_pro.*	AQ490499 RPCI-11-2
27: em_gss_rod.*	AQ216073 HS_3244_B

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283	16.8	449	17	AQ424894
2	183	10.9	382	14	T06700
3	179	10.7	886	13	BF910844
4	174	10.4	1180	14	BQ067695
5	155	9.2	275	12	BF376376
6	116	6.9	364	17	AQ428360

C 80	592	2.5	42	17	B57352	B57352 CIT-HSP-200	153	41	2.4	663	17	AQ543767	AQ543767
C 81	605	2.5	42	17	AQ359848	AQ359848 HS_5027_A	c 154	41	2.4	665	9	AL043052	AL043052
C 82	625	2.5	42	17	AQ377218	AQ377218 RPCI11-14	c 155	41	2.4	667	17	AG105750	AG105750
C 83	642	2.5	42	17	AQ108235	AQ108235 CIT-HSP-2	c 156	41	2.4	672	12	AG498347	AG498347
C 84	663	2.5	42	17	AG081162	AG081162 Pan trogl	c 157	41	2.4	698	17	A2519470	A2519470
C 85	667	2.5	42	17	AG064867	AG064867 Pan trogl	c 158	41	2.4	720	17	AG047945	AG047945
C 86	674	2.5	42	17	AG121331	AG121331 Pan trogl	c 159	41	2.4	725	12	BF237460	BF237460
C 87	685	2.5	42	17	AQ015017	AQ015017 CIT-HSP-2	c 160	41	2.4	736	14	BW178735	BW178735
C 88	687	2.5	42	17	AG180563	AG180563 Pan trogl	c 161	41	2.4	775	14	BQ432755	BQ432755
C 89	737	2.5	42	17	AG085631	AG085631 Pan trogl	c 162	41	2.4	809	17	AQ740956	AQ740956
C 90	786	2.5	42	17	AQ0780489	AQ0780489 HS_3113_A	c 163	41	2.4	821	14	BQ588559	BQ588559
C 91	934	2.5	42	17	AQ232212	AQ232212 AGENCOURT	c 164	41	2.4	849	12	BG393321	BG393321
C 92	972	2.5	42	12	BE068666	BE068666 601446085	c 165	41	2.4	877	10	BE543127	BE543127
C 93	1080	2.5	42	14	BQ917326	BQ917326 AGENCOURT	c 166	41	2.4	901	12	BF247355	BF247355
C 94	197	9	AA322052	AA322052	EST24628	AA322052	c 167	41	2.4	907	14	BQ883658	BQ883658
C 95	206	17	AQ267642	AQ267642	RPCI11-74	AQ267642	c 168	41	2.4	925	12	BF346007	BF346007
C 96	274	9	AA062884	AA062884	z471e03.s	AA062884	c 169	41	2.4	930	12	BF974514	BF974514
C 97	274	9	AA083136	AA083136	zn08a03.r	AA083136	c 170	41	2.4	1055	14	BM810426	BM810426
C 98	307	10	AW832971	AW832971	RC3-TT000	AW832971	c 171	41	2.4	1076	14	BQ711762	BQ711762
C 99	312	17	AQ553837	AQ553837	RPCI-11-3	AQ553837	c 172	41	2.4	1108	12	BF569692	BF569692
C 100	320	10	BE044607	BE044607	hc87c06.x	BE044607	c 173	41	2.4	1271	13	BM458997	BM458997
C 101	334	14	H81700	H81700	ys67a12.sl	H81700	c 174	40	2.4	156	9	AJ230815	AJ230815
C 102	337	10	AW832986	AW832986	RC3-TT000	AW832986	c 175	40	2.4	180	9	AI538252	AI538252
C 103	346	10	BE067459	BE067459	MR4-BT035	BE067459	c 176	40	2.4	204	13	BI016458	BI016458
C 104	357	9	AA05549	AA05549	zw39f03.r	AA05549	c 177	40	2.4	227	10	BE061652	BE061652
C 105	371	17	AQ387288	AQ387288	RPCI11-14	AQ387288	c 178	40	2.4	229	17	AQ424147	AQ424147
C 106	373	14	BQ343483	BQ343483	IL3-NT010	BQ343483	c 179	40	2.4	234	10	AW080836	AW080836
C 107	380	12	BF805872	BF805872	CM1-CI001	BF805872	c 180	40	2.4	245	9	AI540587	AI540587
C 108	386	9	AA829413	AA829413	OQ27D01.s	AA829413	c 181	40	2.4	251	12	BF799220	BF799220
C 109	387	14	H68379	H68379	yr89b09.r1	H68379	c 182	40	2.4	252	9	AA664924	AA664924
C 110	400	9	AA827960	AA827960	OD70B09.s	AA827960	c 183	40	2.4	263	17	AI865776	AI865776
C 111	408	9	AA632581	AA632581	np82B01.s	AA632581	c 184	40	2.4	270	9	AA078381	AA078381
C 112	409	17	AQ215465	AQ215465	zr94h11.r	AQ215465	c 185	40	2.4	278	9	AA731691	AA731691
C 113	409	17	AQ317207	AQ317207	RPCI11-10	AQ317207	c 186	40	2.4	280	17	AQ564761	AQ564761
C 114	416	12	BF808558	BF808558	CM1-CI009	BF808558	c 187	40	2.4	292	9	AA425283	AA425283
C 115	446	9	AA021561	AA021561	z6e9b08.s	AA021561	c 188	40	2.4	302	9	AA665504	AA665504
C 116	450	10	AW304971	AW304971	xv88a07.x	AW304971	c 189	40	2.4	314	17	AQ067606	AQ067606
C 117	451	9	AA722287	AA722287	ag85e05.r	AA722287	c 190	40	2.4	320	12	BF771774	BF771774
C 118	452	10	AW849711	AW849711	IL3-CT021	AW849711	c 191	40	2.4	322	10	AW188912	AW188912
C 119	458	17	B85210	B85210	RPCI11-3014	B85210	c 192	40	2.4	324	9	AI278388	AI278388
C 120	459	17	AQ762370	AQ762370	HS_3190_A	AQ762370	c 193	40	2.4	325	9	AI439786	AI439786
C 121	461	17	AQ286762	AQ286762	RPCI11-89	AQ286762	c 194	40	2.4	342	9	AA411337	AA411337
C 122	461	2.4	461	10	AW860459	AW860459 QV0-CT038	c 195	40	2.4	345	10	AW731858	AW731858
C 123	466	9	AA056250	AA056250	zif62c06.s	AA056250	c 196	40	2.4	350	14	N57059	N57059
C 124	483	10	BE336766	BE336766	ba98d07.y	BE336766	c 197	40	2.4	356	13	BG983308	BG983308
C 125	489	17	AQ242839	AQ242839	HS_2061_A	AQ242839	c 198	40	2.4	359	17	AQ102805	AQ102805
C 126	493	9	AI978782	AI978782	wr59f05.x	AI978782	c 199	40	2.4	371	9	AA909816	AA909816
C 127	503	10	AV754702	AV754702	AV754702	AV754702	c 200	40	2.4	376	9	AI190247	AI190247
C 128	515	17	AQ636787	AQ636787	RPCI-11-4	AQ636787	c 201	40	2.4	380	17	AV683406	AV683406
C 129	523	17	AQ150550	AQ150550	HS_3203_A	AQ150550	c 202	40	2.4	383	9	AA877992	AA877992
C 130	528	14	BM716597	BM716597	UI-E-DX1-	BM716597	c 203	40	2.4	392	9	AA580251	AA580251
C 131	530	9	AA456964	AA456964	aa90f12.s	AA456964	c 204	40	2.4	393	9	AA084212	AA084212
C 132	533	10	AW838204	AW838204	QV2-LT005	AW838204	c 205	40	2.4	396	14	F23221	F23221
C 133	537	17	AQ309780	AQ309780	CITBI-EI-	AQ309780	c 206	40	2.4	399	17	AQ262241	AQ262241
C 134	541	9	AL713027	AL713027	DKF2p686p	AL713027	c 207	40	2.4	401	9	AA293809	AA293809
C 135	547	12	BF736226	BF736226	PM4-KT000	BF736226	c 208	40	2.4	404	12	BF834812	BF834812
C 136	552	17	AQ426785	AQ426785	CITBI-EI-	AQ426785	c 209	40	2.4	409	10	AW243884	AW243884
C 137	557	9	AL589348	AL589348	DKF2p451K	AL589348	c 210	40	2.4	413	9	AI955184	AI955184
C 138	560	17	AQ34909	AQ34909	601458950	AQ34909	c 211	40	2.4	414	12	BE785134	BE785134
C 139	572	12	BM032707	BM032707	K-ESt0100	BM032707	c 212	40	2.4	416	14	R38408	R38408
C 140	576	14	AW827707	AW827707	HS_3118_A	AW827707	c 213	40	2.4	423	17	H94410	H94410
C 141	578	17	AQ779261	AQ779261	RC3-NN006	AQ779261	c 214	40	2.4	423	17	AQ179089	AQ179089
C 142	579	12	BE699328	BE699328	RPCI-11-4	BE699328	c 215	40	2.4	424	12	BF870094	BF870094
C 143	588	17	AQ630833	AQ630833	AGENCOURT	AQ630833	c 216	40	2.4	428	12	BG250390	BG250390
C 144	595	13	BM473240	BM473240	AGENCOURT	BM473240	c 217	40	2.4	429	17	AQ261566	AQ261566
C 145	619	10	BE566446	BE566446	RC3-CI025	BE566446	c 218	40	2.4	433	10	AW157128	AW157128
C 146	622	10	AW854485	AW854485	Pan trogl	AW854485	c 219	40	2.4	437	9	AI038029	AI038029
C 147	640	17	AG083306	AG083306	RPCI-11-3	AG083306	c 220	40	2.4	438	12	BF765657	BF765657
C 148	644	17	AQ528567	AQ528567	Pan trogl	AQ528567	c 221	40	2.4	440	9	AA070899	AA070899
C 149	645	17	AG122818	AG122818	Pan trogl	AG122818	c 222	40	2.4	453	9	AA070345	AA070345
C 150	654	17	AG053415	AG053415	602504026	AG053415	c 223	40	2.4				
C 151	654	17	AG053415	AG053415		AG053415	c 224	40	2.4				
C 152	657	12	BG483444	BG483444		BG483444	c 225	40	2.4				





C 372	38	2.3	494	9	AA160467	AA160467	zq42ei2.r	445	37	2.2	451	17	AQ513895	AQ513895	HS_5152_A
C 373	38	2.3	547	17	AQ539130	AQ539130	RPCI-11-3	c 446	37	2.2	456	17	AQ533235	AQ533235	RPCI-11-3
C 374	38	2.3	548	9	AA553381	AA553381	nk83h05.s	447	37	2.2	461	17	AQ584417	AQ584417	RPCI-11-4
C 375	38	2.3	562	9	AL044286	AL044286	DKF2p343M	448	37	2.2	467	17	AQ126345	AQ126345	HS_3037_B
C 376	38	2.3	588	17	AQ265331	AQ265331	CITBI-E1-	c 449	37	2.2	469	10	AQ1667744	AQ1667744	MR0-HT016
C 377	38	2.3	623	9	AL038936	AL038936	DKF2p566M	450	37	2.2	475	10	AW129434	AW129434	xe15c01.x
C 378	38	2.3	630	13	BI820374	BI820374	603036260	451	37	2.2	478	17	AQ568052	AQ568052	HS_2135_A
C 379	38	2.3	645	17	AQ936487	AQ936487	RPCI-11-S	c 452	37	2.2	482	9	AA164867	AA164867	zq41h02.r
C 380	38	2.3	674	17	AQ485406	AQ485406	RPCI-11-2	c 453	37	2.2	482	10	AV754771	AV754771	AV754771
C 381	38	2.3	689	17	AQ167499	AQ167499	Pan trogl	c 454	37	2.2	488	12	BF997061	BF997061	QV3-GN020
C 382	38	2.3	867	14	BQ430114	BQ430114	AGENCOURT	c 455	37	2.2	489	13	BI013353	BI013353	PM3-ET015
C 383	38	2.3	911	12	BG616384	BG616384	602644450	c 456	37	2.2	492	17	AZ757574	AZ757574	ew09h05.r
C 384	38	2.3	1159	13	BM465227	BM465227	AGENCOURT	c 457	37	2.2	493	17	AQ154511	AQ154511	HS_3014_B
C 385	37	2.2	162	17	AQ427307	AQ427307	CITBI-E1-	c 458	37	2.2	494	14	R13548	R13548	yf59h01.r1
C 386	37	2.2	175	14	BQ358253	BQ358253	QV4-HT026	c 459	37	2.2	495	17	AQ355554	AQ355554	CITBI-E1-
C 387	37	2.2	202	17	AQ008068	AQ008068	RPCI11-20	c 460	37	2.2	495	17	AQ637042	AQ637042	RPCI-11-4
C 388	37	2.2	232	14	BQ357013	BQ357013	RC1-HT022	c 461	37	2.2	499	17	AQ611549	AQ611549	HS_5084_B
C 389	37	2.2	242	13	BI038213	BI038213	IL3-WT028	c 462	37	2.2	509	9	AL037945	AL037945	DKF2p564N
C 390	37	2.2	270	17	BH140963	BH140963	UP_465-F	c 463	37	2.2	510	17	AQ731933	AQ731933	HS_5548_A
C 391	37	2.2	274	17	BH634727	BH634727	UP_422-14	c 464	37	2.2	512	12	BG003985	BG003985	MR3-GN018
C 392	37	2.2	285	13	BG953492	BG953492	CM4-CT062	c 465	37	2.2	513	17	AQ228495	AQ228495	HS_2023_B
C 393	37	2.2	295	9	AL698411	AL698411	DKF2p686I	c 466	37	2.2	526	17	AQ520254	AQ520254	HS_5208_A
C 394	37	2.2	295	17	AQ776038	AQ776038	HS_5570_A	c 467	37	2.2	528	9	AL704544	AL704544	DKF2p686G
C 395	37	2.2	297	9	AA728873	AA728873	nv37b10.i	c 468	37	2.2	528	17	AQ459556	AQ459556	HS_5085_A
C 396	37	2.2	297	9	AA594169	AA594169	nm31g07.s	c 469	37	2.2	533	17	AQ368442	AQ368442	HS_5039_B
C 397	37	2.2	300	10	AW856362	AW856362	RC6-HT028	c 470	37	2.2	535	9	AL704523	AL704523	DKF2p686E
C 398	37	2.2	302	10	AW883708	AW883708	QV2-OT006	c 471	37	2.2	536	17	AQ314955	AQ314955	RPCI-11-2
C 399	37	2.2	305	9	AA478308	AA478308	zu45e10.s	c 472	37	2.2	536	17	AQ415831	AQ415831	UQ1-HF-BN0
C 400	37	2.2	313	9	AA829565	AA829565	od97a11.s	c 473	37	2.2	538	10	AW576374	AW576374	Pan trogl
C 401	37	2.2	316	17	AQ637252	AQ637252	RPCI-11-4	c 474	37	2.2	545	10	AW939214	AW939214	QV1-DT007
C 402	37	2.2	318	10	BE141507	BE141507	MR0-WT008	c 475	37	2.2	549	17	AQ395652	AQ395652	CITBI-E1-
C 403	37	2.2	319	12	BG055729	BG055729	naf87f01.	c 476	37	2.2	559	17	AQ514238	AQ514238	HS_5184_B
C 404	37	2.2	321	9	AA433499	AA433499	zx47d09.s	c 477	37	2.2	572	9	AL700379	AL700379	DKF2p686J
C 405	37	2.2	326	17	AQ372638	AQ372638	RPCI11-15	c 478	37	2.2	574	17	AQ555305	AQ555305	HS_5230_B
C 406	37	2.2	327	14	T07682	T07682	EST05572.Fe	c 479	37	2.2	585	17	AQ528540	AQ528540	RPCI-11-3
C 407	37	2.2	332	12	BF129140	BF129140	601811407	c 480	37	2.2	588	10	AW761491	AW761491	AV761491
C 408	37	2.2	332	9	AA584748	AA584748	no04a11.s	c 481	37	2.2	591	9	AL602276	AL602276	DKF2p686B
C 409	37	2.2	342	9	AA523718	AA523718	n150d10.s	c 482	37	2.2	596	17	AQ581279	AQ581279	RPCI-11-4
C 410	37	2.2	344	9	AA303007	AA303007	EST114046	c 483	37	2.2	607	9	AL110391	AL110391	DKF2p434F
C 411	37	2.2	346	12	BG32847	BG32847	602496318	c 484	37	2.2	607	10	AW377625	AW377625	QV2-CT022
C 412	37	2.2	348	17	B65829	B65829	CIT-HSP-202	c 485	37	2.2	613	12	BG428818	BG428818	602494603
C 413	37	2.2	348	17	AQ478843	AQ478843	RPCI-11-2	c 486	37	2.2	637	10	AW763155	AW763155	Pan trogl
C 414	37	2.2	371	9	AA555229	AA555229	nk32h10.s	c 487	37	2.2	638	10	AV763151	AV763151	AV763151
C 415	37	2.2	372	17	AQ264926	AQ264926	CITBI-E1-	c 488	37	2.2	643	17	AQ088870	AQ088870	Pan trogl
C 416	37	2.2	372	12	BF858318	BF858318	FT020	c 489	37	2.2	661	17	AQ386371	AQ386371	RPCI11-15
C 417	37	2.2	375	10	AW575012	AW575012	UI-HF-BN0	c 490	37	2.2	669	17	AQ088870	AQ088870	Pan trogl
C 418	37	2.2	378	9	AI473838	AI473838	ta17609.x	c 491	37	2.2	670	17	AG135650	AG135650	Pan trogl
C 419	37	2.2	385	9	AA703698	AA703698	ag81c01.r	c 492	37	2.2	673	17	AG163673	AG163673	Pan trogl
C 420	37	2.2	385	10	BE183496	BE183496	CM2-HT067	c 493	37	2.2	675	9	AL699893	AL699893	DKF2p686D
C 421	37	2.2	391	14	T58567	T58567	y62c04.r1	c 494	37	2.2	679	17	AG056576	AG056576	Pan trogl
C 422	37	2.2	392	17	AQ389519	AQ389519	RPCI11-15	c 495	37	2.2	680	9	AL043217	AL043217	DKF2p434I
C 423	37	2.2	399	9	AL700720	AL700720	DKF2p686B	c 496	37	2.2	687	17	AG012282	AG012282	Homo sapi
C 424	37	2.2	405	9	AI281881	AI281881	qt68h12.x	c 497	37	2.2	687	17	AG093469	AG093469	Pan trogl
C 425	37	2.2	405	9	AA129000	AA129000	zo08g10.r	c 498	37	2.2	696	17	AG095345	AG095345	Pan trogl
C 426	37	2.2	407	17	AQ393742	AQ393742	CITBI-E1-	c 499	37	2.2	707	17	AQ423788	AQ423788	CITBI-E1-
C 427	37	2.2	407	17	AQ547512	AQ547512	RPCI-11-4	c 500	37	2.2	717	17	AG183361	AG183361	Pan trogl
C 428	37	2.2	407	17	AQ618012	AQ618012	HS_5165_B	c 501	37	2.2	717	17	AQ0392510	AQ0392510	CITBI-E1-
C 429	37	2.2	413	9	AI671695	AI671695	ly26h02.x	c 502	37	2.2	721	17	AG183712	AG183712	Pan trogl
C 430	37	2.2	414	17	AG115980	AG115980	RPCI11-56	c 503	37	2.2	722	17	AG013254	AG013254	Pan trogl
C 431	37	2.2	417	14	BQ214493	BQ214493	AGENCOURT	c 504	37	2.2	742	17	AG013265	AG013265	Homo sapi
C 432	37	2.2	418	17	AG193568	AG193568	CIT-HSP-2	c 505	37	2.2	742	17	AG029645	AG029645	602388811
C 433	37	2.2	426	13	BI024249	BI024249	CM3-MT038	c 506	37	2.2	743	12	BG290544	BG290544	CIT-HSP-2
C 434	37	2.2	428	17	AZ520869	AZ520869	RPCI-11-2	c 507	37	2.2	759	17	AQ110904	AQ110904	CIT-HSP-2
C 435	37	2.2	431	17	AQ498911	AQ498911	HS_5154_A	c 508	37	2.2					
C 436	37	2.2	432	10	AW074933	AW074933	xa66c01.x	c 509	37	2.2					
C 437	37	2.2	432	17	AG117054	AG117054	HS_2200_A	c 510	37	2.2					
C 438	37	2.2	433	17	AQ143045	AQ143045	HS_3064_A	c 511	37	2.2					
C 439	37	2.2	434	9	AI126262	AI126262	qc55d12.x	c 512	37	2.2					
C 440	37	2.2	435	9	AI239509	AI239509	qh42b08.x	c 513	37	2.2					
C 441	37	2.2	438	9	AA478368	AA478368	zu45e10.r	c 514	37	2.2					
C 442	37	2.2	438	17	AQ753413	AQ753413	HS_5310_B	c 515	37	2.2					
C 443	37	2.2	451	12	BG012219	BG012219	RC2-GN026	c 516	37	2.2					
C 444	37	2.2	451	17	AQ242020	AQ242020	CITBI-E1-	c 517	37	2.2					

c 518	37	2.2	795	13	B1917148	B1917148	603181568	c 591	36	2.1	471	14	T59509	T59509 yb65b12.sl	
519	37	2.2	893	14	BQ706862	AGENCOURT	BQ706862	AGENCOURT	592	36	2.1	475	17	AQ582744	RPCI-11-4
520	37	2.2	909	14	BQ411407	AGENCOURT	BQ411407	AGENCOURT	593	36	2.1	479	13	BG955867	PM3-CT064
521	37	2.2	924	12	BG426043	602492723	BG426043	602492723	594	36	2.1	480	9	AI814682	wk66b07.x
522	37	2.2	940	14	BQ933539	AGENCOURT	BQ933539	AGENCOURT	595	36	2.1	483	14	BM708254	UI-E-C11-
523	37	2.2	950	17	AQ747336	HS_5536_A	AQ747336	HS_5536_A	596	36	2.1	485	17	AQ380313	RPCI-11-1
524	37	2.2	956	13	B762107	603049285	B762107	603049285	597	36	2.1	487	9	AI291439	qm73f01.x
525	36	2.1	126	17	AQ634965	RPCI-11-4	AQ634965	RPCI-11-4	598	36	2.1	487	14	N48297	yy77e12.sl
526	36	2.1	134	14	H97666		H97666	ym07c01.sl	599	36	2.1	492	17	AQ622080	HS_5302_B
527	36	2.1	156	13	BG993409	MR3-HT099	BG993409	MR3-HT099	600	36	2.1	495	12	BG054544	7o45h12.x
528	36	2.1	157	10	AW071109	wv87d09.x	AW071109	wv87d09.x	601	36	2.1	495	17	AQ461423	HS_5202_A
529	36	2.1	182	13	BG953902	RC4-CT068	BG953902	RC4-CT068	602	36	2.1	498	9	AL699780	DRF2p6861
530	36	2.1	192	9	AA502428	ng62a04.s	AA502428	ng62a04.s	603	36	2.1	499	9	AI819964	wj88g07.x
531	36	2.1	203	12	BG152927	nah94f07.	BG152927	nah94f07.	604	36	2.1	502	9	AI761357	w160g12.x
532	36	2.1	248	12	BF853596	MR2-EN009	BF853596	MR2-EN009	605	36	2.1	507	9	AI271323	q103a03.x
533	36	2.1	265	9	AI626133	ar90g12.x	AI626133	ar90g12.x	606	36	2.1	509	17	AQ377149	RPCI-11-14
534	36	2.1	273	10	AW062454	MRO-CT006	AW062454	MRO-CT006	607	36	2.1	510	17	AQ541941	RPCI-11-3
535	36	2.1	296	9	AA297930	EST11351	AA297930	EST11351	608	36	2.1	513	17	AQ675818	HS_2145_A
536	36	2.1	308	10	AW731785		AW731785	bc41g08.x	609	36	2.1	516	10	BE645220	7e66f01.x
537	36	2.1	311	10	AW933182		AW933182	rc2-DT000	610	36	2.1	517	14	T55992	yb36a05.sl
538	36	2.1	352	9	AA350186		AA350186	EST57392	611	36	2.1	520	17	A2520256	RPCI-11-7
539	36	2.1	323	9	AI251590	qv45c09.x	AI251590	qv45c09.x	612	36	2.1	520	17	AQ675767	HS_2145_A
540	36	2.1	328	9	AA502688	ng63a04.s	AA502688	ng63a04.s	613	36	2.1	533	10	AW937040	PM2-DT003
541	36	2.1	328	17	AQ200403	RPCI11-63	AQ200403	RPCI11-63	614	36	2.1	545	9	AU147829	AU147829
542	36	2.1	338	9	AA083374		AA083374	zn09g01.s	615	36	2.1	549	17	AQ785945	HS_3053_A
543	36	2.1	352	9	AA083333	zn09g01.r	AA083333	zn09g01.r	616	36	2.1	552	10	AV721509	AV721509
544	36	2.1	363	17	AQ242287	HS_2063_B	AQ242287	HS_2063_B	617	36	2.1	554	9	AU158674	AU158674
545	36	2.1	368	12	BF933433	CM1-NT027	BF933433	CM1-NT027	618	36	2.1	556	10	AW836820	CV1-LT003
546	36	2.1	368	17	AQ221468	HS_3138_B	AQ221468	HS_3138_B	619	36	2.1	561	9	AL697074	DRF2p6860
547	36	2.1	368	17	AQ586276		AQ586276	RPCI-11-4	620	36	2.1	563	17	AQ759224	HS_2206_B
548	36	2.1	375	9	AA364707		AA364707	EST75397	621	36	2.1	565	17	AQ482585	RPCI-11-2
549	36	2.1	377	17	B79675		B79675	CIT-HSP-204	622	36	2.1	568	10	AW971855	EST383944
550	36	2.1	378	17	AQ209451		AQ209451	HS_3234_A	623	36	2.1	581	10	BE258837	601117378
551	36	2.1	379	9	AA528520		AA528520	nf01c03.s	624	36	2.1	589	17	AQ323381	RPCI11-10
552	36	2.1	382	17	B79296		B79296	CIT978SK-12	625	36	2.1	590	10	AV760725	AV760725
553	36	2.1	384	9	AA834633		AA834633	cd98b04.s	626	36	2.1	593	14	BQ019757	UI-E-ED0-
554	36	2.1	388	9	AI349961		AI349961	ca97h01.x	627	36	2.1	600	9	AL597249	DRF2p313A
555	36	2.1	391	9	AA262832		AA262832	zs24c04.r	628	36	2.1	602	9	AI064781	HA0554_Hu
556	36	2.1	393	9	AI401007		AI401007	tg91e09.x	629	36	2.1	602	12	BG778393	602666556
557	36	2.1	397	12	BF762725		BF762725	CM0-CS004	630	36	2.1	607	12	BG778393	602666556
558	36	2.1	398	9	AA593478		AA593478	nn26h01.s	631	36	2.1	608	12	BE879668	601491858
559	36	2.1	400	13	BG952075		BG952075	MR1-CT073	632	36	2.1	612	14	BM697717	UI-E-DX0-
560	36	2.1	402	9	AI130709		AI130709	qc12d09.x	633	36	2.1	612	14	BQ777571	1138b04.x
561	36	2.1	407	17	AQ146849	HS_2244_A	AQ146849	HS_2244_A	634	36	2.1	613	12	BF920000	MR1-NT017
562	36	2.1	403	9	AI469450	LM13a06.x	AI469450	LM13a06.x	635	36	2.1	616	12	BF920137	AGENCOURT
563	36	2.1	404	17	AQ825034		AQ825034	HS_5248_A	636	36	2.1	616	14	BQ894378	MR1-NT017
564	36	2.1	405	9	AI536858		AI536858	to13h05.x	637	36	2.1	621	10	AW976904	EST389013
565	36	2.1	405	10	AW852961		AW852961	RC1-CT024	638	36	2.1	629	12	BF245970	601864779
566	36	2.1	409	9	AI537395		AI537395	tp39g12.x	639	36	2.1	631	17	AG102586	Pan trogl
567	36	2.1	414	9	AA807684		AA807684	nv66g10.s	640	36	2.1	631	17	AG112015	Pan trogl
568	36	2.1	415	12	BF763357		BF763357	CM1-CS003	641	36	2.1	631	17	AQ484812	RPCI-11-2
569	36	2.1	415	13	BG942940		BG942940	ax31e03.x	642	36	2.1	634	17	AG048715	Pan trogl
570	36	2.1	415	17	AQ117895		AQ117895	HS_3000_B	643	36	2.1	639	17	B55278	CIT-HSP-386
571	36	2.1	421	9	AL045917		AL045917	DKFp434L	644	36	2.1	641	13	B1038954	IL3-NT028
572	36	2.1	422	17	AQ204181		AQ204181	HS_3107_A	645	36	2.1	642	17	AG149348	Pan trogl
573	36	2.1	423	9	AA309530		AA309530	EST180398	646	36	2.1	644	12	BF970596	602273909
574	36	2.1	427	10	AW272815		AW272815	xu24d06.x	647	36	2.1	647	17	AG049716	Pan trogl
575	36	2.1	430	9	AI345173		AI345173	tb80h02.x	648	36	2.1	648	10	AV760012	AV760012
576	36	2.1	430	10	AW78780		AW78780	ho12h11.x	649	36	2.1	648	17	AG0609217	HS_3455_A
577	36	2.1	430	17	AQ828069		AQ828069	HS_5311_B	650	36	2.1	651	12	BE866206	601679004
578	36	2.1	433	9	AI160144		AI160144	qb63b09.x	651	36	2.1	653	14	BQ635924	hd02a11.y
579	36	2.1	434	17	B35251		B35251	HS-1027-A2-	652	36	2.1	664	13	BG939617	cr59b10.x
580	36	2.1	437	9	AL597532		AL597532	DKF2p313G	653	36	2.1	667	17	AG135237	Pan trogl
581	36	2.1	437	9	AA484718		AA484718	ne65b05.s	654	36	2.1	674	14	BM997701	UI-H-DH0-
582	36	2.1	438	10	AW698965		AW698965	EST381943	655	36	2.1	677	17	AG080153	Pan trogl
583	36	2.1	438	17	AQ021410		AQ021410	CIT-HSP-2	656	36	2.1	681	10	AV707244	AV707244
584	36	2.1	441	17	B71035		B71035	CIT-HSP-206	657	36	2.1	688	17	AG103337	Pan trogl
585	36	2.1	444	17	AQ109317		AQ109317	CIT-HSP-3	658	36	2.1	692	17	AG100650	Pan trogl
586	36	2.1	456	17	AQ819671		AQ819671	HS_5326_B	659	36	2.1	693	12	BG167219	602345362
587	36	2.1	458	14	N47820		N47820	vw95b07.r1	660	36	2.1	704	13	BM007670	603617095
588	36	2.1	461	17	AQ629782	RPCI-11-1	AQ629782	RPCI-11-1	661	36	2.1	709	12	BG562791	602581611
589	36	2.1	465	10	AW148775	xf04a05.x	AW148775	xf04a05.x	662	36	2.1	713	10	AW190277	x113c06.x
590	36	2.1	470	14	W94202	zelle10.sl	W94202	zelle10.sl	663	36	2.1	717	12	BF823689	MR1-RT003



810	35	2.1	410	17	AQ336567 HS_5006_A	AQ336567 HS_5006_A	c 883	35	2.1	500	13	BM054883
811	35	2.1	411	9	A1924076 wn78e05.x	A1924076 wn78e05.x	c 884	35	2.1	501	9	AI821400
812	35	2.1	411	17	AQ474283 CITBI-EI-	AQ474283 CITBI-EI-	c 885	35	2.1	502	9	AI692471
813	35	2.1	413	9	A1335387 tb79c10.x	A1335387 tb79c10.x	c 886	35	2.1	503	17	AQ885036
814	35	2.1	415	9	AA493594 nh01b03.s	AA493594 nh01b03.s	c 887	35	2.1	503	17	AQ437756
815	35	2.1	415	17	AQ091738 HS_3015_A	AQ091738 HS_3015_A	c 888	35	2.1	507	17	AQ497131
816	35	2.1	416	17	AQ553557 RPCI-11-4	AQ553557 RPCI-11-4	c 889	35	2.1	512	10	AW851405
817	35	2.1	417	9	AI284126 qt72g12.x	AI284126 qt72g12.x	c 890	35	2.1	512	17	AQ318544
818	35	2.1	419	9	AA625565 zu90f03.s	AA625565 zu90f03.s	c 891	35	2.1	516	17	AQ165095
819	35	2.1	420	9	AA489944 ab01h06.r	AA489944 ab01h06.r	c 892	35	2.1	517	17	AQ137785
820	35	2.1	422	9	AA206036 zq77e06.s	AA206036 zq77e06.s	c 893	35	2.1	518	17	AQ351567
821	35	2.1	423	17	AQ091777 HS_3015_A	AQ091777 HS_3015_A	c 894	35	2.1	519	17	AQ442067
822	35	2.1	428	9	AI820561 ye12d04.y	AI820561 ye12d04.y	c 895	35	2.1	522	17	AQ485045
823	35	2.1	430	17	AQ471340 CITBI-EI-	AQ471340 CITBI-EI-	c 896	35	2.1	523	17	AQ573002
824	35	2.1	434	17	AQ319040 RPCI-11-10	AQ319040 RPCI-11-10	c 897	35	2.1	524	10	BE246758
825	35	2.1	435	9	AA054158 zfs5b08.s	AA054158 zfs5b08.s	c 898	35	2.1	525	17	AA565660
826	35	2.1	439	9	AA515728 ng70f04.s	AA515728 ng70f04.s	c 899	35	2.1	525	17	AQ630634
827	35	2.1	440	9	AA099564 z177g11.s	AA099564 z177g11.s	c 900	35	2.1	526	10	BE301088
828	35	2.1	444	13	BI008396 MR4-RF002	BI008396 MR4-RF002	c 901	35	2.1	526	12	BF111319
829	35	2.1	445	14	BQ023054 UT-1-BB1P	BQ023054 UT-1-BB1P	c 902	35	2.1	526	17	B54504
830	35	2.1	447	17	AQ338601 HS_2202_B	AQ338601 HS_2202_B	c 903	35	2.1	527	17	B63808
831	35	2.1	448	17	AQ402993 HS_5053_A	AQ402993 HS_5053_A	c 904	35	2.1	530	9	AI142551
832	35	2.1	449	9	AI151490 qc74h06.x	AI151490 qc74h06.x	c 905	35	2.1	530	17	AQ486009
833	35	2.1	449	17	B42790 HS-1056-B1-	B42790 HS-1056-B1-	c 906	35	2.1	531	17	AQ684857
834	35	2.1	450	10	AW081436 xc31a06.x	AW081436 xc31a06.x	c 907	35	2.1	531	17	AQ233139
835	35	2.1	450	12	BF950207 QV0-NN114	BF950207 QV0-NN114	c 908	35	2.1	532	12	BF964838
836	35	2.1	451	9	AI653493 qt95a01.x	AI653493 qt95a01.x	c 909	35	2.1	534	17	AQ282060
837	35	2.1	453	14	N63991 yz81g08.s1	N63991 yz81g08.s1	c 910	35	2.1	535	17	AQ282440
838	35	2.1	454	14	R02571 ye80d08.s1	R02571 ye80d08.s1	c 911	35	2.1	536	17	AQ386773
839	35	2.1	455	17	AQ244920 HS_3204_B	AQ244920 HS_3204_B	c 912	35	2.1	536	17	AQ410364
840	35	2.1	456	14	D44625 HUMSPUP023	D44625 HUMSPUP023	c 913	35	2.1	537	9	AI158290
841	35	2.1	456	14	R81361 yf04g07.s1	R81361 yf04g07.s1	c 914	35	2.1	537	14	BM767212
842	35	2.1	459	9	AA114102 zn66g09.s	AA114102 zn66g09.s	c 915	35	2.1	537	10	AW684881
843	35	2.1	459	10	AW117723 xe34g12.x	AW117723 xe34g12.x	c 916	35	2.1	547	10	AW971160
844	35	2.1	460	10	AW117704 xs34e12.x	AW117704 xs34e12.x	c 917	35	2.1	547	17	AQ483873
845	35	2.1	461	10	AW165741 MR3-GN034	AW165741 MR3-GN034	c 918	35	2.1	548	9	AI145083
846	35	2.1	463	13	BI054057 MR3-GN034	BI054057 MR3-GN034	c 919	35	2.1	549	9	AI179985
847	35	2.1	464	14	BM688445 UT-E-CLO-	BM688445 UT-E-CLO-	c 920	35	2.1	554	17	AQ451966
848	35	2.1	465	13	BI043686 PM0-OT020	BI043686 PM0-OT020	c 921	35	2.1	556	17	AQ678294
849	35	2.1	465	17	AQ079792 CIT-HSP-2	AQ079792 CIT-HSP-2	c 922	35	2.1	558	17	AZ518496
850	35	2.1	467	10	BE069194 QV3-BT037	BE069194 QV3-BT037	c 923	35	2.1	559	17	AQ276157
851	35	2.1	467	13	BI054800 MR3-GN034	BI054800 MR3-GN034	c 924	35	2.1	560	17	AQ568715
852	35	2.1	467	17	AQ242954 HS_2061_A	AQ242954 HS_2061_A	c 925	35	2.1	566	12	BF213410
853	35	2.1	469	9	AI904840 RC-BT068-	AI904840 RC-BT068-	c 926	35	2.1	568	12	BF213410
854	35	2.1	469	17	B99076 CIT-HSP-228	B99076 CIT-HSP-228	c 927	35	2.1	569	12	AW935783
855	35	2.1	470	10	BE156471 QV0-HT036	BE156471 QV0-HT036	c 928	35	2.1	569	12	BM163916
856	35	2.1	470	12	BF770328 RC1-IT001	BF770328 RC1-IT001	c 929	35	2.1	570	17	AG033642
857	35	2.1	471	9	AA493475 nh04a09.s	AA493475 nh04a09.s	c 930	35	2.1	570	17	AQ540562
858	35	2.1	471	9	AA526424 n160c01.s	AA526424 n160c01.s	c 931	35	2.1	573	14	BM992855
859	35	2.1	471	17	B58212 CIT-HSP-201	B58212 CIT-HSP-201	c 932	35	2.1	576	9	AI145134
860	35	2.1	472	14	N49881 yv24a05.s1	N49881 yv24a05.s1	c 933	35	2.1	584	14	BQ181732
861	35	2.1	473	17	AQ225556 HS_2012_A	AQ225556 HS_2012_A	c 934	35	2.1	587	17	AG125461
862	35	2.1	475	10	AV719213 AV719213	AV719213 AV719213	c 935	35	2.1	590	17	AQ313925
863	35	2.1	476	10	AW973722 EST385823	AW973722 EST385823	c 936	35	2.1	590	17	AQ412636
864	35	2.1	478	12	EG054978 nnc93a04.	EG054978 nnc93a04.	c 937	35	2.1	591	12	BG722651
865	35	2.1	479	9	AI338185 qt48f12.x	AI338185 qt48f12.x	c 938	35	2.1	592	14	BM675461
866	35	2.1	482	17	AQ112301 CIT-HSP-2	AQ112301 CIT-HSP-2	c 939	35	2.1	602	10	AW979292
867	35	2.1	483	10	AW367815 MR0-HT016	AW367815 MR0-HT016	c 940	35	2.1	602	17	AQ538479
868	35	2.1	484	10	AW272281 xul7b03.x	AW272281 xul7b03.x	c 941	35	2.1	602	17	AQ540860
869	35	2.1	484	10	AW965597 EST377670	AW965597 EST377670	c 942	35	2.1	603	17	AQ386438
870	35	2.1	484	14	BM673259 UT-E-CR1-	BM673259 UT-E-CR1-	c 943	35	2.1	605	17	AQ535932
871	35	2.1	484	14	BM697760 UT-E-DX0-	BM697760 UT-E-DX0-	c 944	35	2.1	610	17	AQ584602
872	35	2.1	484	17	AQ128409 HS_3080_A	AQ128409 HS_3080_A	c 945	35	2.1	610	17	AQ388943
873	35	2.1	484	17	AQ237976 RPCI-11-65	AQ237976 RPCI-11-65	c 946	35	2.1	614	14	D56387
874	35	2.1	485	14	BQ328080 MR0-BT003	BQ328080 MR0-BT003	c 947	35	2.1	614	17	AQ627870
875	35	2.1	490	10	BE069130 QV3-BT037	BE069130 QV3-BT037	c 948	35	2.1	617	10	AW859036
876	35	2.1	493	13	AQ218748 HS_3248_B	AQ218748 HS_3248_B	c 949	35	2.1	621	17	AG046735
877	35	2.1	494	12	BF943868 QV0-NN114	BF943868 QV0-NN114	c 950	35	2.1	624	17	AQ384737
878	35	2.1	495	9	AI921613 wo27f11.x	AI921613 wo27f11.x	c 951	35	2.1	626	17	AG146372
879	35	2.1	496	9	AI032037 ow68b06.x	AI032037 ow68b06.x	c 952	35	2.1	627	9	AA130054
880	35	2.1	498	13	BI857941 603388625	BI857941 603388625	c 953	35	2.1	628	17	AQ263554
881	35	2.1	499	14	BQ023003 UT-1-BB1P	BQ023003 UT-1-BB1P	c 954	35	2.1	635	17	AZ519410
882	35	2.1	500	10	BE080715 QV1-BT063	BE080715 QV1-BT063	c 955	35	2.1	635	17	AQ239846

```
956 35 2.1 643 17 AG088874 Pan trogl
957 35 2.1 652 17 A0527391 RPTI-11-3
958 35 2.1 654 17 A0197090 CIT-HSP-2
959 35 2.1 655 14 A0762854 AV762854
960 35 2.1 655 14 BQ183715 UI-H-EUO-
961 35 2.1 656 17 AG059884 Pan trogl
962 35 2.1 658 17 AG160832 Pan trogl
963 35 2.1 658 9 AL118628
964 35 2.1 658 14 BM718434 UI-E-E01-
965 35 2.1 660 17 AG040423 Pan trogl
966 35 2.1 662 17 AG104573 Pan trogl
967 35 2.1 662 17 A0528678 RPTI-11-3
968 35 2.1 664 17 A0625527 CITBI-E1-
969 35 2.1 665 17 A2521084 RPTI-11-2
970 35 2.1 666 17 A0349547 RPTI-11-11
971 35 2.1 671 17 AG131672 Pan trogl
972 35 2.1 675 10 AW979240
973 35 2.1 679 17 B68278
974 35 2.1 681 17 AG115075
975 35 2.1 682 17 AG169815
976 35 2.1 684 17 AG185247
977 35 2.1 684 17 A0632073 RPTI-11-4
978 35 2.1 685 17 AG159422
979 35 2.1 687 14 BM989960
980 35 2.1 690 17 A0054102 CIT-HSP-2
981 35 2.1 692 17 AG100946
982 35 2.1 693 9 AL048528
983 35 2.1 695 9 AL704309
984 35 2.1 696 17 AG169302
985 35 2.1 703 17 AG171907
986 35 2.1 704 17 AG175213
987 35 2.1 704 17 AG048625 Pan trogl
988 35 2.1 706 17 AG067699 Pan trogl
989 35 2.1 707 17 A0540349 RPTI-11-3
990 35 2.1 711 12 BE439676
991 35 2.1 721 12 BG695931
992 35 2.1 723 10 BE258830
993 35 2.1 723 17 AG172328
994 35 2.1 727 14 BQ000199
995 35 2.1 731 12 BG485697
996 35 2.1 733 10 AW963594
997 35 2.1 740 17 AG109194
998 35 2.1 742 17 A0538005 RPTI-11-3
999 35 2.1 743 14 BQ774932
1000 35 2.1 743 17 AG168820 Pan trogl
```

## ALIGNMENTS

```
RESULT 1
A0424894 449 bp DNA linear GSS 24-MAR-1999
LOCUS CITBI-E1-2576P11.TF CITBI-E1 Homo sapiens genomic clone 2576P11,
DEFINITION DNA sequence.
ACCESSION A0424894
VERSION A0424894.1 GI:4498160
KEYWORDS GSS.
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 449)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CITBI-E1-2576P11.TR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
```

FEATURES  
source

```
1..449
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2576P11"
/clone_lib="CITBI-E1"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
CalTech Human BAC Library D"
BASE COUNT 134 a 98 c 123 g 94 t
ORIGIN
```

```
Query Match 16.8%; Score 283; DB 17; Length 449;
Best Local Similarity 99.7%; Pred. No. 1.8e-44;
Matches 333; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 37 GGTCTCTACAAAAATACAAAATTAGCTGGATTGGTGGCGGTGCTCATGCCCTATAATC 96
|||||
Db 34 GGTCTCTACAAAAATACAAAATTAGCTGGATTGGTGGCGGTGCTCATGCCCTATAATC 93
|||||
QY 97 CCAGCACTTTGGGAGCCTGAGTGGTGGATCCTCAAGTCAGGAGTTCAGACTAGCC 156
|||||
Db 94 CCAGCACTTTGGGAGCCTGAGTGGTGGATCCTCAAGTCAGGAGTTCAGACTAGCC 153
|||||
QY 157 TGGCCACATGGTGAACCCCTATCTCTACTGAAATAYAAAAAGCTAGAGTGGTGGCAC 216
|||||
Db 154 TGGCCACATGGTGAACCCCTATCTCTACTGAAATACAAAAGCTAGAGTGGTGGCAC 213
|||||
QY 217 ACACCTGTAATCCAGCTACTTAGGAGCCTCAGCAGGAGAGTAATGCTTGAAGCCTAGAGG 276
|||||
Db 214 ACACCTGTAATCCAGCTACTTAGGAGCCTCAGCAGGAGAGTAATGCTTGAAGCCTAGAGG 273
|||||
QY 277 TGAAGTTGTAGTGAGCCGAGATTGTCATCATTTGCACAAATGGAGGGAGCCACCAGCCTGG 336
|||||
Db 274 TGAAGTTGTAGTGAGCCGAGATTGTCATCATTTGCACAAATGGAGGGAGCCACCAGCCTGG 333
|||||
QY 337 GCACACAGAGGAATCTCCGCTCCCAAAAAAAA 370
|||||
Db 334 GCACACAGAGGAATCTCCGCTCCCAAAAAAAA 367
```

```
RESULT 2
T06700/c 382 bp mRNA linear EST 30-JUN-1993
LOCUS EST04589 Fetal brain, Stratagene (cat#936206) Homo sapiens cDNA
DEFINITION clone HFBDX16 similar to EST containing Alu repeat, mRNA sequence.
ACCESSION T06700
VERSION T06700.1 GI:317849
KEYWORDS EST.
SOURCE human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 382)
AUTHORS Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.
TITLE 3,400 expressed sequence tags identify diversity of transcripts
from human brain
JOURNAL Nat. Genet. 4, 256-267 (1993)
COMMENT MEDLINE 93364420
Contact: Adams, MD
The Institute for Genomic Research
932 Clopper Road, Gaithersburg, MD 20878
Tel: 3018699056
Fax: 3018699423
```

Email: mdadams@tigr.org  
Seq primer: M13-21.  
Location/Qualifiers  
source  
1. .382  
/organism="Homo sapiens"  
/db\_xref="ATCC (inhost):83354"  
/db\_xref="taxon:9606"  
/clone="HFBDX16"  
/clone\_lib="Fetal brain, Stratagene (cat#936206)"  
/note="Vector: LambdaZAP-II; 17-18 wk gestation, female;  
oligo-dT + random primed cDNA synthesis; lambdaZAP-II  
vector, 1.0kb average inser size."  
BASE COUNT 76 a 101 c 90 g 111 t 4 others  
ORIGIN

/clone\_lib="NIH\_MGC\_118"  
/tissue\_type="leukocyte"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV  
(destroyed); RNA source leukocytes from anonymous pool of  
non-activated adult donors. Library is oligo-dT primed  
and directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.7 kb, insert size range  
1.2-3.3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 027. Note:  
this is a NIH\_MGC Library."  
BASE COUNT 255 a 244 c 224 g 163 t  
ORIGIN

Query Match 10.9%; Score 183; DB 14; Length 382;  
Best Local Similarity 99.3%; Pred. No. 1e-25; Indels 0; Gaps 0;  
Matches 283; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 54 CAAATTTAGCTGGGATGGGTGGCTCAGCCCTATATCCAGCACCTTTGGGAGCC 113  
|||||  
Db 307 CAAATTTAGCTGGGATGGGTGGCTCAGCCCTATATCCAGCACCTTTGGGAGCC 248  
|||||  
Qy 114 TGAGTGGGTGGATCAGCTGAAGTCAGAGTTCAGACTAGCTGGCCACATGGTGA 173  
|||||  
Db 247 TGAGTGGGTGGATCAGCTGAAGTCAGAGTTCAGACTAGCTGGCCACATGGTGA 188  
|||||  
Qy 174 CCTTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGGCACACACTGTGTAATCCCGAGC 233  
|||||  
Db 187 CCTTATCTCTACTGAAATACAAAAGCTAGACGTGGTGGCACACACTGTGTAATCCCGAGC 128  
|||||  
Qy 234 TACTTAGAGGCTGAGCGAGGAGAAATTCCTTGAAGCCCTAGAGGTGAAGTTGTAGTGAGC 293  
|||||  
Db 127 CACTTAGAGGCTGAGCGAGGAGAAATTCCTTGAAGCCCTAGAGGTGAAGTTGTAGTGAGC 68  
|||||  
Qy 294 CGAGATTGCATCTGACCAATGAGGGAGCCACCGCTGGGC 338  
|||||  
Db 67 CGAGATTGCATCTGACCAATGAGGGAGCCACCGCTGGGC 23  
|||||

RESULT 3  
BI910844/c  
LOCUS  
DEFINITION BI910844 886 bp mRNA linear EST 16-OCT-2001  
60306887F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5217977 5',  
mRNA sequence.  
ACCESSION BI910844  
VERSION BI910844.1 GI:16174292  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 886)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-f@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL1547 row: m column: 18  
High quality sequence start: 4  
High quality sequence stop: 432.

FEATURES  
source  
1. .886  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5217977"

Query Match 10.7%; Score 179; DB 13; Length 886;  
Best Local Similarity 99.6%; Pred. No. 2.5e-25;  
Matches 229; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 517 AAAAAAATAAAGCAACATATCTCGAACAAGGATCTCTCATAGCTTCCCAACAGAT 576  
|||||  
Db 312 AAAAAAATAAAGCAACATATCTCGAACAAGGATCTCTCATAGCTTCCCAACAGAT 253  
|||||  
Qy 577 TTCTAATCAGAAACATGAGGCCAGAAAGCAGTGGAGGAGGACRACCTTCAGGAGGCCG 636  
|||||  
Db 252 TTCTAATCAGAAACATGAGGCCAGAAAGCAGTGGAGGAGGACRACCTTCAGGAGGCCG 193  
|||||  
Qy 637 GGAGGATGTTGTACAGGCTGGGGCAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGA 696  
|||||  
Db 192 GGAGGATGTTGTACAGGCTGGGGCAAGGGCTTCCGGCTACCAACTGGGAGCTCTGGGA 133  
|||||  
Qy 697 ACAGCCCTGTTGCAACAAGAGCCATAGCCGGCCAGAGCCAGGAATG 746  
|||||  
Db 132 ACAGCCCTGTTGCAACAAGAGCCATAGCCGGCCAGAGCCAGGAATG 83  
|||||

RESULT 4  
BQ067695  
LOCUS  
DEFINITION BQ067695 1180 bp mRNA linear EST 02-APR-2002  
AGENCOURT\_6643098 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5768406  
5', mRNA sequence.  
ACCESSION BQ067695  
VERSION BQ067695.1 GI:19896741  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1180)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-f@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLNL12828 row: d column: 07  
High quality sequence start: 3  
High quality sequence stop: 536.

FEATURES  
source  
1. .1180  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5768406"  
/clone\_lib="NIH\_MGC\_121"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source anonymous pool of 3







Matches 166; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY 197	AAAGCTAGACGTGGTGGCACACACCTGTATATCCAGCTACTTAGGAGGCTGAGCGAGGAG 256
Db	
	194 AAAGCTAGACGTGGTGGCACACACCTGTATATCCAGCTACTTAGGAGGCTGAGCGAGGAG 253
QY 257	AATTGCTTGAAGCCCTAGAGGTGAAGTTGTAGTAGGAGCCAGATTGCATCATTTGCACAATG 316
Db	
	254 AATTGCTTGAAGCCCTAGAGGTGAAGTTGTAGTAGGAGCCAGATTGCATCATTTGCACAATG 313
QY 317	GAGGGAGCCACCGAGCTGGGCAACAAGAGGAAATCTCGCTCTCCAA 363
Db	
	314 GAGGGAGCCACCGAGCTGGGCAACAAGAGGAAATCTCGCTCTCCAA 360
RESULT 7	
LOCUS	BI030898 258 bp mRNA linear EST 14-JUN-2001
DEFINITION	IL5-MT0262-290301-406-c07 MT0262 Homo sapiens cDNA, mRNA sequence.
ACCESSION	BI030898
VERSION	BI030898.1 GI:14437528
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	1 (bases 1 to 258)
	Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
	Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
	Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
	,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
	Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags	
	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL	20202663
MEDLINE	Contact: Simpson A.J.G.
COMMENT	Laboratory of Cancer Genetics
	Ludwig Institute for Cancer Research
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
	Brazil
	Tel: +55-11-2704922
	Fax: +55-11-2707001
	Email: asimpson@ludwig.org.br
	This sequence was derived from the FAPESP/LICR Human Cancer Genome
	Project. This entry can be seen in the following URL
	(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-MT0262-
	290301-406-c07&t3=2001-03-29&t4=1)
	Seq primer: puc 18 forward
	High quality sequence stop: 255.
FEATURES	Location/Qualifiers
	1..258
source	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone_lib="MT0262"
	/dev_stage="Adult"
	/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
	SmaI; A mini-library was made by cloning products derived
	from ORESTES PCR (U.S. Letters Patent application No. 196
	,716 - Ludwig Institute for Cancer Research) profiles
	into the puc 18 vector. Reverse transcription of tissue
	mRNA and cDNA amplification were performed under low
	stringency conditions."
	80 a 63 c 68 g 47 t
BASE COUNT	
ORIGIN	
Query Match 3.2%; Score 53; DB 13; Length 258;	
Best Local Similarity 100.0%; Pred. No. 0.3;	
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 125	GATCACCCTGAAGTCAGGAGTTCAGACTAGCCTGGCCAAACATGGTGAACCCCT 177

Db 101	GATCACCCTGAAGTCAGGAGTTCAGACTAGCCTGGCCAAACATGGTGAACCCCT 153
RESULT 8	
LOCUS	BF746029 304 bp mRNA linear EST 10-JAN-2001
DEFINITION	CM4-BT0858-071100-423-g01 BT0858 Homo sapiens cDNA, mRNA sequence.
ACCESSION	BF746029
VERSION	BF746029.1 GI:12072705
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
AUTHORS	1 (bases 1 to 304)
	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
	Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
	Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
	Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
	,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
	Simpson,A.J.
	Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags	
	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL	20202663
MEDLINE	Contact: Simpson A.J.G.
COMMENT	Laboratory of Cancer Genetics
	Ludwig Institute for Cancer Research
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
	Brazil
	Tel: +55-11-2704922
	Fax: +55-11-2707001
	Email: asimpson@ludwig.org.br
	This sequence was derived from the FAPESP/LICR Human Cancer Genome
	Project. This entry can be seen in the following URL
	(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-BT0858-
	071100-423-g01&t3=2000-11-07&t4=1)
	Seq primer: puc 18 forward
	High quality sequence start: 2
	High quality sequence stop: 302.
	Location/Qualifiers
FEATURES	1..304
	/organism="Homo sapiens"
source	/db_xref="taxon:9606"
	/clone_lib="BT0858"
	/dev_stage="Adult"
	/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
	SmaI; A mini-library was made by cloning products derived
	from ORESTES PCR (U.S. Letters Patent application No. 196
	,716 - Ludwig Institute for Cancer Research) profiles
	into the puc 18 vector. Reverse transcription of tissue
	mRNA and cDNA amplification were performed under low
	stringency conditions."
	91 a 78 c 81 g 53 t 1 others
BASE COUNT	
ORIGIN	
Query Match 3.2%; Score 53; DB 12; Length 304;	
Best Local Similarity 100.0%; Pred. No. 0.26;	
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 125	GATCACCCTGAAGTCAGGAGTTCAGACTAGCCTGGCCAAACATGGTGAACCCCT 177
Db 169	GATCACCCTGAAGTCAGGAGTTCAGACTAGCCTGGCCAAACATGGTGAACCCCT 221
RESULT 9	
LOCUS	BF910677 327 bp mRNA linear EST 18-JAN-2001
DEFINITION	RC1-UT0033-021100-018-f12 UT0033 Homo sapiens cDNA, mRNA sequence.
ACCESSION	BF910677
VERSION	BF910677.1 GI:12302135
KEYWORDS	EST.
SOURCE	human.

ORGANISM	Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 327)	
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.	
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	
MEDLINE	20202663	
COMMENT	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC1&t2=RC1-UT0033-021100-018-f12&t3=2000-11-02&t4=1) Seq primer: puc 18 forward High quality sequence start: 8 High quality sequence stop: 327.	
FEATURES	location/Qualifiers	
source	1..327	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone_lib="UT0033"	
	/dev_stage="Adult"	
	/note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."	
BASE COUNT	110 a 64 c 87 g 66 t	
ORIGIN		
Query Match	3.1%; Score 52; DB 12; Length 327;	
Best Local Similarity	100.0%; Pred. No. 0.37;	
Matches	52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	208 TGGTGGCACACCTGTATCCAGCTACTTAGGAGGCTGAGCGAGAGAAT 259	
Db	45 TGGTGGCACACCTGTATCCAGCTACTTAGGAGGCTGAGCGAGAGAAT 96	
RESULT 10		
BF913166		
LOCUS	RC1-UT0033-241100-110-f10 UT0033 Homo sapiens cDNA, mRNA sequence.	
DEFINITION		
ACCESSION	BF913166	
VERSION	BF913166.1 GI:12304728	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 327)	
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.	
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed	

sequence tags		
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		
20202663		
Contact: Simpson A.J.G.		
Laboratory of Cancer Genetics		
Ludwig Institute for Cancer Research		
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil		
Tel: +55-11-2704922		
Fax: +55-11-2707001		
Email: asimpson@ludwig.org.br		
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL		
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC1&t2=RC1-UT0033-241100-110-f10&t3=2000-11-24&t4=1)		
Seq primer: puc 18 forward		
High quality sequence stop: 222.		
location/Qualifiers		
1..327		
/organism="Homo sapiens"		
/db_xref="taxon:9606"		
/clone_lib="UT0033"		
/dev_stage="Adult"		
/note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."		
BASE COUNT	112 a 63 c 85 g 67 t	
ORIGIN		
Query Match	3.1%; Score 52; DB 12; Length 327;	
Best Local Similarity	100.0%; Pred. No. 0.37;	
Matches	52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	208 TGGTGGCACACCTGTATCCAGCTACTTAGGAGGCTGAGCGAGAGAAT 259	
Db	41 TGGTGGCACACCTGTATCCAGCTACTTAGGAGGCTGAGCGAGAGAAT 92	
RESULT 11		
B81880		
LOCUS	B81880 528 bp DNA linear GSS 09-APR-1999	
DEFINITION	RPC111-1706.TP RPC1-11 Homo sapiens genomic clone RPC1-11-1706, DNA sequence.	
ACCESSION	B81880	
VERSION	B81880.1 GI:2868903	
KEYWORDS	GSS.	
SOURCE	human.	
ORGANISM	Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 528)	
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.	
TITLE	Use of BAC End Sequences for Sequence-Ready Map Building (1998)	
JOURNAL	Unpublished (1998)	
COMMENT	Other_GSSs: RPC111-1706.TVB Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdamads@tigr.org Clones are derived from the human BAC library RPC1-11. For BAC library availability, please contact pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page:	

http://www.tigr.org/tdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: SP6  
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..528

/organism="Homo sapiens"

/db\_xref="GDB:7506485"

/db\_xref="taxon:9606"

/clone="RPCI-11-1706"

/clone\_lib="RPCI-11"

/sex="Male"

/cell\_type="Lymphocytes"

/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;

RPC111 Human Male BAC Library"

169 a 111 c 109 g 139 t

BASE COUNT

ORIGIN

Query Match 3.0%; Score 51; DB 17; Length 528;

Best Local Similarity 100.0%; Pred. No. 0.35;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 217 ACACCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGAAATGCTTGAA 267

|||||

Db 59 ACACCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGAAATGCTTGAA 109

RESULT 12

AQ352156

LOCUS

DEFINITION C17BI-E1-2537J20.TF C17BI-E1 Homo sapiens genomic clone 2537J20,

DNA sequence.

ACCESSION AQ352156

VERSION AQ352156.1 GI:4179491

KEYWORDS GSS.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 206)

AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and

Venter,J.C.

TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready

Map Building

JOURNAL Unpublished (1997)

COMMENT Other\_GSS: C17BI-E1-2537J20.TR

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

http://www.tigr.org/tdb/hungen/bac\_end\_search/bac\_end\_search.html.

Seq primer: M13-21

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..206

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="2537J20"

/clone\_lib="C17BI-E1"

/sex="male"

/cell\_type="sperm"

/note="Vector: pBeloBAC11; Site\_1: EcoRI; Site\_2: EcoRI;

Caltech Human BAC Library D"

63 a 44 c 58 g 41 t

BASE COUNT

ORIGIN

Query Match 3.0%; Score 50; DB 17; Length 206;

Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 218 CACCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGAAATGCTTGAA 267

|||||

Db 138 CACCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGAAATGCTTGAA 187

RESULT 13

AQ405838

LOCUS

DEFINITION

HS\_5043\_B2\_B04\_T7 RPCI-11 Human Male BAC Library Homo sapiens

genomic clone Plate-619 Col-8 Row-D, DNA sequence.

ACCESSION AQ405838

VERSION AQ405838.1 GI:4415826

KEYWORDS GSS.

SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 519)

AUTHORS Mahairas,G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and

scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999);

MEDLINE 99380589

COMMENT

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm)

or from Research h Genetics (info@resgen.com). BAC end Web Server:

http://www.htsc.washington.edu

Plate: 619 row: D column: 8

Seq primer: T7

Class: BAC ends

High quality sequence stop: 519.

FEATURES

source

Location/Qualifiers

1..519

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="Plate-619 Col-8 Row-D"

/clone\_lib="RPCI-11 Human Male BAC Library"

/sex="male"

/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;

Male blood DNA was isolated from one randomly chosen donor

and partially digested with a combination of EcoRI and

EcoRI Methylase. Size selected DNA was cloned into the

pBACE3.6 vector at EcoRI sites"

BASE COUNT 154 a 122 c 122 g 117 t

ORIGIN

Query Match 3.0%; Score 50; DB 17; Length 519;

Best Local Similarity 100.0%; Pred. No. 0.55;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 218 CACCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGAAATGCTTGAA 267

|||||

Db 414 CACCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGAAATGCTTGAA 463

RESULT 14

AQ309829

LOCUS

DEFINITION

C17-HSP-2385N1.TRB C17-HSP Homo sapiens genomic clone 2385N1, DNA

sequence.

ACCESSION AQ309829

```

VERSION      AQ309829.1  GI:4041716
KEYWORDS     GSS.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
              1 (bases 1 to 559)
              Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
              Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
              Venter,J.C.
TITLE        Use of a random human BAC End Sequence Database for Sequence-Ready
JOURNAL      Map Building
COMMENT      Unpublished (1998)
              Contact: Shaying Zhao, William Nierman, Mark Adams
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: hbe@tigr.org
              Clones are available from Research Genetics (info@resgen.com). BAC
              end search page:
              http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
              Seq primer: M13 Reverse
              Class: BAC ends.

FEATURES     Location/Qualifiers
             1..559
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="2385N1"
             /clone_lib="CIT-HSP"
             /sex="Male"
             /cell_type="Sperm"
             /note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
             HindIII"
BASE COUNT   172 a 111 c 127 g 149 t
ORIGIN
Query Match      3.0%; Score 50; DB 17; Length 559;
Best Local Similarity 100.0%; Pred. No. 0.51;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  218  CACCTGTATCCAGCTACTTAGGAGGCTGAGCAGGAGAGTAATGCTTGAA 267
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  458  CACCTGTATCCAGCTACTTAGGAGGCTGAGCAGGAGAGTAATGCTTGAA 507

RESULT 15
AQ563791/c
LOCUS      AQ563791          584 bp    DNA    linear    GSS 29-MAY-1999
DEFINITION HS_5338_B1_B05_77A RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate-914 Col-9 Row=D, DNA sequence.
ACCESSION  AQ563791
VERSION    AQ563791.1  GI:4923262
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 584)
AUTHORS    Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE      Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE    99380589
COMMENT    Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887

```

```

Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 914 row: D column: 9
Seq primer: F7
Class: BAC ends
High quality sequence stop: 584.
Location/Qualifiers
1..584
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate-914 Col-9 Row=D"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"
BASE COUNT   136 a 131 c 102 g 205 t
ORIGIN
Query Match      2.9%; Score 49; DB 17; Length 584;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  219  ACCTGTATCCAGCTACTTAGGAGGCTGAGCAGGAGAGTAATGCTTGAA 267
      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db  370  ACCTGTATCCAGCTACTTAGGAGGCTGAGCAGGAGAGTAATGCTTGAA 322

RESULT 16
AA484228
LOCUS      AA484228          310 bp    mRNA    linear    EST 14-AUG-1997
DEFINITION ne99d07.s1 NCI-CGAP_Kidl Homo sapiens cDNA clone IMAGE:911437
            similar to contains Alu repetitive element;contains element MER22
            repetitive element ;, mRNA sequence.
ACCESSION  AA484228
VERSION    AA484228.1  GI:2213041
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 310)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapsb-r@mail.nih.gov
            Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 431 Std Error: 0.00
            Seq primer: -41ml3 fwd.ET from Amersham.
            Location/Qualifiers
            1..310
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:911437"
            /clone_lib="NCI-CGAP_Kidl"
            /tissue_type="kidney"
            /lab_host="DH10B"

```

/note="Vector: pAMP10; mRNA made from invasive kidney tumor, cDNA made by oligo-dt priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."

BASE COUNT 90 a 75 c 88 g 57 t

Query Match 2.9%; Score 48; DB 9; Length 310;  
Best Local Similarity 100.0%; Pred. No. 2.2;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAATTCCTTCAA 267  
|||||

Db 168 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAATTCCTTCAA 215

RESULT 17  
LOCUS BF677661 361 bp mRNA linear EST 21-DEC-2000  
DEFINITION 602085543F1 NIH\_MGC\_83 Homo sapiens cDNA clone IMAGE:4249715 5',  
mRNA sequence.

ACCESSION BF677661

VERSION BF677661.1 GI:11951556

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 361)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: CLONETECH Laboratories, Inc.

cDNA Library Preparation: CLONETECH Laboratories, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone Sequencing by: Incyte Genomics, Inc.

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1CM1070 row: e column: 12

High quality sequence stop: 361.

Location/Qualifiers

1..361

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="IMAGE:4249715"

/lab\_host="NIH\_MGC\_83"

/note="Organ: prostate; Vector: pDNR-LIB (Clontech);

Site\_1: Sfil (ggcgctcgcc); Site\_2: Sfil (ggccattatggc

); 5' and 3' adaptors were used in cloning as follows: 5'

adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor

sequence: 5'-ATTCTAGAGCCGAGCGGCGCATG-dt(30)BN-3'

(where B = A, C, G, or T). Average

insert size 1.4 kb (range 0.5-4.0 kb). 14/15 colonies

contained inserts by PCR. This library was enriched for

full-length clones and was constructed by Clontech

Laboratories (Palo Alto, CA)."

BASE COUNT 119 a 76 c 97 g 69 t

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 2.9%; Score 48; DB 12; Length 361;

Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAATTCCTTCAA 267

|||||

Db 203 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAATTCCTTCAA 250

RESULT 18

LOCUS AQ469676/c

DEFINITION

CITBI-EI-2587123.TR CITBI-EI Homo sapiens genomic clone 2587123,

DNA sequence.

ACCESSION AQ469676

VERSION AQ469676.1 GI:4653330

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 404)

AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and

Venter,J.C.

TITLE Use of BAC End Sequences from CalTech Libraries for Sequence-Ready

Map Building

JOURNAL Unpublished (1997)

COMMENT Other\_GSSs: CITBI-EI-2587123.TF

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: [hbe@tigr.org](mailto:hbe@tigr.org)

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

[http://www.tigr.org/tldb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html).

Seq primer: M13 Reverse

Class: BAC ends.

Location/Qualifiers

1..404

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_image="2587123"

/clone\_lib="CITBI-EI"

/sex="male"

/cell\_type="sperm"

/note="Vector: pBeloBAC11; Site\_1: EcoRI; Site\_2: EcoRI;

Caltech Human BAC Library D"

BASE COUNT 100 a 103 c 77 g 123 t

ORIGIN

Query Match 2.9%; Score 48; DB 17; Length 404;

Best Local Similarity 100.0%; Pred. No. 1.7;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAATTCCTTCAA 267

|||||

Db 263 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAATTCCTTCAA 216

|||||

RESULT 19

LOCUS BF771528

DEFINITION

IL5-IT0027-291100-293-g02 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF771528

VERSION BF771528.1 GI:12119428

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 489)

AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL  
MEDLINE  
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-IT0027-  
291100-293-g02&t3=2000-11-29&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 447.

## FEATURES

source

1. .489  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IT0027"  
/dev\_stage="Adult"  
/note="Organ: epid.tumor; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 89 a 146 c 115 g 137 t 2 others  
ORIGIN

Query Match 2.9%; Score 48; DB 12; Length 489;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGATTGCTTGA 267  
|||||  
Db 351 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGATTGCTTGA 304  
|||||

RESULT 20  
A0395268/c  
LOCUS

DEFINITION  
A0395268  
VERSION  
A0395268.1 GI:4366294  
KEYWORDS  
GSS.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 519)  
AUTHORS  
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and  
Venter,J.C.  
TITLE  
Use of BAC End Sequences from CalTech Libraries for Sequence-Ready  
Map Building  
JOURNAL  
Unpublished (1997)  
COMMENT  
Other GSSs: CITBI-E1-2546L13.TF  
Contact: Shaving Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: M13 Reverse  
Class: BAC ends.  
FEATURES  
Location/Qualifiers

source

1. .519  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="2546L13"  
/clone\_lib="CITBI-E1"  
/sex="male"  
/cell\_type="sperm"  
/note="Vector: pBelOBAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
CalTech Human BAC Library D"  
BASE COUNT 157 a 107 c 103 g 151 t 1 others  
ORIGIN

Query Match 2.9%; Score 48; DB 17; Length 519;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGATTGCTTGA 267  
|||||  
Db 451 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGATTGCTTGA 404  
|||||

## RESULT 21

AW971963

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```

SOURCE          human.
ORGANISM        Homo sapiens
REFERENCE       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE          Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
               Venter,J.C.
JOURNAL        Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
COMMENT        Map Building
               Unpublished (1997)
               Other GSSs: CITBI-E1-2556D7.TR
               Contact: Shaying Zhao, William Nierman, Mark Adams
               Department of Eukaryotic Genomics
               The Institute for Genomic Research
               9712 Medical Center Dr., Rockville, MD 20850
               Tel: 301 838 0200
               Fax: 301 838 0208
               Email: hbe@tigr.org
               Clones are available from Research Genetics (info@resgen.com). BAC
               end search page:
               http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
               Seq primer: M13-21
               Class: BAC ends.
FEATURES       Location/Qualifiers
               source
                 1..719
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
                 /clone="2556D7"
                 /clone_lib="CITBI-E1"
                 /sex="male"
                 /cell_type="sperm"
                 /note="vector: pheloBAC11; site_1: EcoRI; site_2: EcoRI;
                 CalTech Human BAC Library D"
BASE COUNT     277 a 115 c 135 g 190 t      2 others
ORIGIN
Query Match    2.9%; Score 48; DB 17; Length 719;
Best Local Similarity 100.0%; Pred. No. 0.94;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  220  CCTGTAATCCCGAGCTACTTAGGAGGCTGAGGAGGAGAGAAATTCCTTGAA 267
      |||||||
DB   83  CCTGTAATCCCGAGCTACTTAGGAGGCTGAGGAGGAGAGAAATTCCTTGAA 130

RESULT 23
LOCUS      AG057132              735 bp      DNA      linear      GSS 02-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-043K01.R, genomic survey sequence.
ACCESSION  AG057132
VERSION    AG057132.1 GI:16594591
KEYWORDS   GSS.
SOURCE     Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
           BAC Library clone:PTB-043K01.R.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE  1
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
           Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of Library PTB
JOURNAL    Unpublished
AUTHORS    2 (bases 1 to 735)
           Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
           Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL    Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
           and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
           1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
           (E-mail:chimbesc@gscc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
           Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB this BAC end
           was generated during the R&D process and may have higher chance of
           clone tracking errors.
PRIMERS
Sequencing: TJ
LIBRARY
Vector       : pBACE3.6
R.Site 1     : EcoRI
R.Site 2     : EcoRI.
Location/Qualifiers
               source
                 1..738
                 /organism="Pan troglodytes"
                 /db_xref="taxon:9598"
                 /clone="RP43-023104.TJ"
                 /sex="male"
                 /cell_type="lymphocytes"
BASE COUNT     222 a 159 c 169 g 186 t      2 others
ORIGIN

```

```

clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector       : pKSL45
R.Site 1     : SacI
R.Site 2     : SacI.
Location/Qualifiers
               source
                 1..735
                 /organism="Pan troglodytes"
                 /db_xref="taxon:9598"
                 /clone="PTB-043K01.R"
                 /sex="male"
                 /cell_type="lymphoblast"
                 /clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT     233 a 162 c 151 g 178 t      11 others
ORIGIN
Query Match    2.9%; Score 48; DB 17; Length 735;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  220  CCTGTAATCCCGAGCTACTTAGGAGGCTGAGGAGGAGAGAAATTCCTTGAA 267
      |||||||
DB   382  CCTGTAATCCCGAGCTACTTAGGAGGCTGAGGAGGAGAGAAATTCCTTGAA 429

RESULT 24
LOCUS      AG158320              738 bp      DNA      linear      GSS 09-JAN-2002
DEFINITION Pan troglodytes DNA, clone: RP43-023104.TJ, genomic survey
           sequence.
ACCESSION  AG158320
VERSION    AG158320.1 GI:16687998
KEYWORDS   GSS.
SOURCE     Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
           Male BAC Library clone:RP43-023104.TJ.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE  1
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
           Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of Library RPCI-43
JOURNAL    Unpublished
AUTHORS    2 (bases 1 to 738)
           Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
           Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL    Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
           and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
           1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
           (E-mail:chimbesc@gscc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
           Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
           end was generated during the R&D process and may have higher chance
           of clone tracking errors.
PRIMERS
Sequencing: TJ
LIBRARY
Vector       : pBACE3.6
R.Site 1     : EcoRI
R.Site 2     : EcoRI.
Location/Qualifiers
               source
                 1..738
                 /organism="Pan troglodytes"
                 /db_xref="taxon:9598"
                 /clone="RP43-023104.TJ"
                 /sex="male"
                 /cell_type="lymphocytes"
BASE COUNT     222 a 159 c 169 g 186 t      2 others
ORIGIN

```

```

Query Match      2.9%; Score 48; DB 17; Length 736;
Best Local Similarity 100.0%; Pred. No. 0.91;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGAAATGCTTGAA 267
|||||
Db 366 CCTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAGAAATGCTTGAA 413

RESULT 25
AQ937369
LOCUS      AQ937369      278 bp      DNA      linear      GSS 23-AUG-2000
DEFINITION NBI-875R Human NotI clones Homo sapiens genomic, DNA sequence.
ACCESSION  AQ937369
VERSION     AQ937369.1 GI:7213747
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 278)
AUTHORS   Zabarovsky,E.R., Gizatullin,R., Podowski,R.M., Zabarovska,V.V., Xie
            L., Muravenko,O.V., Kozirev,S., Petrenko,L., Skobeleva,N., Li,D.,
            Protopopov,A., Kashuba,V., Ernberg,I., Winberg,G. and Wahlstedt,C.
            NotI clones in the analysis of the human genome
            Nucleic Acids Res. 28 (7), 1635-1639 (2000)
            20175728
COMMENT    Contact: Podowski RM
            Center for Genomics Research
            Karolinska Institute
            17177 Stockholm, Sweden
            Tel: +46-8-728-6372
            Fax: +46-8-337983
            Email: Raf.Podowski@cgr.ki.se
            Class: NotI site.
FEATURES   source
            Location/Qualifiers
            1..278
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="Human NotI clones"
            /note="organ: Lung; DNA was isolated from A549 cells after
            sodium arsenite exposure for 4 weeks. This fragment was
            differentially methylated relative to untreated controls
            and was identified using methylation sensitive AP-PCR and
            sequenced."
BASE COUNT    79 a 69 c 72 g 56 t 2 others
ORIGIN
Query Match      2.8%; Score 47; DB 17; Length 278;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 CAGGAGTTCAGACATAGCTGCGCAACATGGTGAACCCATCTCTA 184
|||||
Db 116 CAGGAGTTCAGACATAGCTGCGCAACATGGTGAACCCATCTCTA 162

RESULT 26
B92374
LOCUS      B92374      346 bp      DNA      linear      GSS 25-JUN-1998
DEFINITION CIT-HSP-2172G14.TF CIT-HSP Homo sapiens genomic clone 2172G14, DNA
            sequence.
ACCESSION  B92374
VERSION     B92374.1 GI:2974711
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 346)
AUTHORS   Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
            Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,

Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.
FEATURES   Location/Qualifiers
            1..346
            /organism="Homo sapiens"
            /db_xref="GDB:710383"
            /db_xref="taxon:9606"
            /clone_lib="CIT-HSP"
            /sex="Male"
            /cell_type="Sperm"
            /note="Vector: pBelorAC11; Site_1: HindIII; Site_2:
            HindIII"
BASE COUNT    96 a 71 c 88 g 91 t
ORIGIN
Query Match      2.8%; Score 47; DB 17; Length 346;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 CTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAATGCTTGAA 267
|||||
Db 192 CTGTAATCCAGCTACTTAGGAGGCTGAGGAGGAGAAATGCTTGAA 238

RESULT 27
BF918155/c
LOCUS      BF918155      500 bp      mRNA      linear      EST 18-JAN-2001
DEFINITION IL3-UT0115-261200-408-B08 UT0115 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF918155
VERSION     BF918155.1 GI:12309613
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 500)
AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            20202663
COMMENT    Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-UT0115-
            261200-408&t3=2000-12-26&t4=1)

```



Seq primer: puc 18 forward  
High quality sequence stop: 465.

#### FEATURES

Source  
1..500  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="UT0115"  
/dev\_stage="Adult"

/note="Organ: uterus\_tumor; Vector: puc18; Site.1: SmaI; Site.2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 103 a 124 c 124 g 149 t

#### ORIGIN

Query Match 2.8%; Score 47; DB 12; Length 500;

Best Local Similarity 100.0%; Pred. No. 2.1;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 CCTGAAGTCAGGAGTTCAGACTAGCTGGCCCAACATGGTGAACCC 176

Db 292 CCTGAAGTCAGGAGTTCAGACTAGCTGGCCCAACATGGTGAACCC 246

#### RESULT 28

AG099522/c

LOCUS Pan troglodytes DNA, clone: PTB-101N05.R, genomic survey sequence.  
ACCESSION AG099522  
VERSION AG099522.1 GI:16720039

KEYWORDS

SOURCE

Pan troglodytes male lymphoblast DNA, clone\_lib:PTB Chimpanzee Male  
BAC Library clone:PTB-101N05.R.

#### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

#### REFERENCE

AUTHORS

Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,

Totoki,Y., Watanabe,H. and Sakaki,Y.

BAC end sequences of Library PTB

#### TITLE

Unpublished

#### JOURNAL

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@psc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

#### PRIMERS

Sequencing: M13Rev

#### LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI.

#### FEATURES

Source

Location/Qualifiers

1..658

/organism="Pan troglodytes"

/db\_xref="taxon:9598"

/clone\_lib="PTB-101N05.R"

/sex="male"

/cell\_type="lymphoblast"

/clone\_lib="PTB Chimpanzee Male BAC Library"

#### BASE COUNT

#### ORIGIN

129 a 201 c 122 g 204 t

2 others

#### Query Match

Best Local Similarity 100.0%; Pred. No. 1.6;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 CACCTGTATCCAGCTACTTAGGAGGCTGAGCAGGAGGAATTGCTT 264

Db 356 CACCTGTATCCAGCTACTTAGGAGGCTGAGCAGGAGGAATTGCTT 310

#### RESULT 29

AQ937652

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Human NotI clones Homo sapiens genomic, DNA sequence.

1 (bases 1 to 989)

REFERENCE

AUTHORS

Zabarovsky,E.R., Gizatullin,R., Podowski,R.M., Zabarovsky,V.V., Xie

,L., Muravenko,O.V., Kozyrev,S., Petrenko,I., Skobeleva,N., Li,J.,

Protopopov,A., Kashuba,V., Ernberg,I., Winberg,G. and Wahlestedt,C.

NotI clones in the analysis of the human genome

Nucleic Acids Res. 28 (7), 1635-1639 (2000)

20175728

COMMENT

Contact: Podowski RM

Center for Genomics Research

Karolinska Institute

17177 Stockholm, Sweden

Tel: +46-8-728-6372

Fax: +46-8-337983

Email: Raf.Podowski@cgr.ki.se

Class: NotI site.

Location/Qualifiers

1..989

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="Human NotI clones"

/note="Organ: Lung; DNA was isolated from A549 cells after

sodium arsenite exposure for 4 weeks. This fragment was

differentially methylated relative to untreated controls

and was identified using methylation sensitive AP-PCR and

sequenced."

BASE COUNT 225 a 293 c 213 g 256 t

2 others

#### ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 1;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 CAGAGTTCAGACTAGCTGGCCACATGGTGAACCCCTATCTCTA 184

Db 115 CAGAGTTCAGACTAGCTGGCCACATGGTGAACCCCTATCTCTA 161

#### RESULT 30

T47924

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 122)

Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins

T47924 yb18c10.r1 StrataGene fetal spleen (#937205) Homo sapiens cDNA  
clone IMAGE:71538 5' similar to contains TAR1 repetitive element.  
mRNA sequence.

T47924 T47924.1 GI:649904

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 (bases 1 to 122)

Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,

Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins

M., M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevasaks, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)

COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu

```

FEATURES
  source
    high quality sequence stop: 90.
    location/Qualifiers
      1. 122
        /organism="Homo sapiens"
        /db_xref="CDB:493203"
        /db_xref="taxon:9606"
        /clone="IMAGE:71538"
        /clone_lib="Stratagene fetal spleen (#937205)"
        /tissue_type="fetal spleen"
        /dev_stage="fetal"
        /lab_host="SOLR cells (kanamycin resistant)"
        /note="organ: spleen; Vector: pBluescript SK-; Site:1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Pooled spleens. Average insert size: 1.0 kb; Uni-2AP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCACGTTTTTTTTTTTTTTT 3'
          29 a 27 c 37 g 26 t 3 others
BASE COUNT

```

```

RESULT 31
AA971049/c
LOCUS          368 bp      mRNA      linear      EST 13-APR-1999
DEFINITION    op10h01.s1 NCI_CGAP_kid6 Homo Sapiens cDNA clone IMAGE:157313 3'
               similar to contains Alu repetitive element:contains L1.t3 L1
               repetitive element ; , mRNA sequence.

```

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 368)  
REFERENCE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
CONTACT: Robert Strausberg. Ph. D.  
COMMENT

found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 1367 Std Error: 0.00  
 Seq primer: -40mi3 fwd. ET from Amersham  
 High quality sequence stop: 325.

```

source
1. .368
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1575313"
/clone_lib="NCI_CGAP_Kid6"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: kidney; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
oligo dt. Pooled kidney tumors, 5' adaptor sequence: 5'
GAATCGGCACGAG 3' 3' adaptor sequence: 5'
CTCGAGCTTTTTTTTTTTTTTTT 3' Average insert size: 1.0 kb."
BASE COUNT      114 a   80 c   61 g   113 t
ORIGIN

Query Match      2.7%; Score 46; DB 9; Length 368;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218  CACCTGTAATCCCAGCTACTTTAGGAGGCTGAGCAGGAGAAATTGCT 263
        |||||||
Db 172  CACCTGTAATCCCAGCTACTTTAGGAGGCTGAGCAGGAGAAATTGCT 127

RESULT 32
AQ490499
LOCUS      AQ490499
DEFINITION RPCI-11-230012.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-230012
          . DNA sequence.
          GSS 24-APR-1999

```

REFERENCE	1 (bases 1 to 503)
AUTHORS	Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.
TITLE	Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

JOURNAL COMMENT

unpublished (1997)  
Other\_GSSs: RPL-11-230012.TV  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbeetiq.org

```

FEATURES
source
Location/Qualifiers
1..503
/organism="Homo sapiens"
/db_xref="GBD:7588283"
/db_xref="taxon:9606"
/clone="RPC1-11-230012"
/clone_lib="RPC1-11"
/sex="Male"

```

```
/cell_type="Lymphocytes"
/Note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC11 Human Male BAC Library"
BASE COUNT      148 a   102 c   136 g   113 t      4 others
ORIGIN

Query Match      2.7%; Score 46; DB 17; Length 503;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGCTGGAACCCCTATCT 181
|||||
Db 174 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGCTGGAACCCCTATCT 219
|||||

RESULT 33
A0216073/c
LOCUS
DEFINITION
HS_3244_B1_D05_r7 CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3244 Col=9 Row=H, DNA sequence.
ACCESSION
A0216073
VERSION
A0216073.1 GI:3631393
KEYWORDS
GSS.
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 509)
AUTHORS
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE
99380589
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@edjeong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 909 row: L column: 16
Seq primer: T7
Class: BAC ends
High quality sequence stop: 509.
Location/Qualifiers
1..559
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3244 Col=9 Row=H"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/Note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      124 a   122 c   104 g   157 t      2 others
ORIGIN

Query Match      2.7%; Score 46; DB 17; Length 509;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 CTATATCCAGCAGCTTTGGGAGCCTGAGGTGGTGGATCACCTGA 134
|||||
Db 443 CTATATCCAGCAGCTTTGGGAGCCTGAGGTGGTGGATCACCTGA 398
|||||

RESULT 34
A0569744
LOCUS
DEFINITION
HS_5333_B2_F08_r7A RPCI-11 Human Male BAC Library Homo sapiens
genomic clone Plate=909 Col=16 Row=L, DNA sequence.
ACCESSION
A0569744
```

```
VERSION
A0569744.1 GI:4962964
KEYWORDS
GSS.
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 555)
AUTHORS
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE
99380589
COMMENT
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@edjeong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 909 row: L column: 16
Seq primer: T7
Class: BAC ends
High quality sequence stop: 555.
Location/Qualifiers
1..555
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=909 Col=16 Row=L"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/Note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"
BASE COUNT      162 a   116 c   133 g   141 t      3 others
ORIGIN

Query Match      2.7%; Score 46; DB 17; Length 555;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGCTGGAACCCCTATCT 181
|||||
Db 181 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGCTGGAACCCCTATCT 226
|||||

RESULT 35
AG175442/c
LOCUS
DEFINITION
Pan troglodytes DNA, clone: RP43-046G04.T7, genomic survey
sequence.
ACCESSION
AG175442
VERSION
AG175442.1 GI:16705122
KEYWORDS
GSS.
SOURCE
Male BAC Library clone:RP43-046G04.T7.
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1
AUTHORS
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE
BAC end sequences of Library RPCI-43
JOURNAL
Unpublished
```



RESULT 38  
AQ105410/c  
LOCUS  
DEFINITION  
HS\_3023\_B1\_G01\_MR CIT Approved Human Genomic Sperm Library D Homo  
sapiens genomic clone Plate-3023 Col-1 Row=N, DNA sequence.  
ACCESSION  
AQ105410  
VERSION  
AQ105410.1 GI:3480766  
KEYWORDS  
GSS.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 262)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
TITLE  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
MEDLINE  
99380589  
COMMENT  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 3023 row: N column: 1  
Class: BAC ends  
High quality sequence stop: 262.  
Location/Qualifiers  
1..262  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="Plate-3023 Col=1 Row=N"  
/clone\_lib="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in  
E-Coli DH10B"  
BASE COUNT 51 a 76 c 69 g 66 t  
ORIGIN  
Query Match 2.6%; Score 44; DB 17; Length 262;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 207 GTGGTGGCACACCTGTAATCCAGCTACTTAGGAGCGTGAGG 250  
|||||  
Db 116 GTGGTGGCACACACCTGTAATCCAGCTACTTAGGAGCGTGAGG 73  
|||||  
RESULT 39  
AA630672/c  
LOCUS  
DEFINITION  
ae65902.s1 Stratagene lung carcinoma 937218 Homo sapiens cDNA clone  
IMAGE:951794 3' similar to contains Alu repetitive element;; mRNA  
sequence.  
ACCESSION  
AA630672  
VERSION  
AA630672.1 GI:2553283  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 281)  
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
Krizman,D., Kubaca,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin  
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,  
White,Y., Wylie,T., Waterston,R. and Willson,R.  
washU-NCI human EST Project  
TITLE

Unpublished (1997)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
This clone is available royalty-free through LNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40m13 fwd. Et from Amersham  
High quality sequence stop: 241.  
Location/Qualifiers  
1..281  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:951794"  
/clone\_lib="Stratagene lung carcinoma 937218"  
/tissue\_type="lung carcinoma"  
/cell\_line="NCI-H69"  
/dev\_stage="cell line NCI-H69"  
/lab\_host="SOLR (kanamycin resistant)"  
/note="Organ: lung; Vector: pBluescript SK-; Site\_1: EcoRI  
; Site\_2: XhoI; Cloned unidirectionally. Primer: Oligo  
dT. Small cell carcinoma cell line NCI-H69. Average  
insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor  
sequence: 5' GAAATTCGGCAGCAG 3' -3' adaptor sequence: 5'  
CTCGAGTTTCTTTTCTTTTCTTTT 3'  
BASE COUNT 51 a 70 c 65 g 95 t  
ORIGIN  
Query Match 2.6%; Score 44; DB 9; Length 281;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 224 TAATCCAGCTACTTAGGAGCGTGAGGAGAGAAATTCCTGAA 267  
|||||  
Db 147 TAATCCAGCTACTTAGGAGCGTGAGGAGAGAAATTCCTGAA 104  
|||||  
RESULT 40  
BG943962  
LOCUS  
DEFINITION  
ax44d12.x1 Proliferating Human Erythroid Cells (LCB:ax library)  
Homo sapiens cDNA clone ax44d12 random, mRNA sequence.  
ACCESSION  
BG943962  
VERSION  
BG943962.1 GI:14343334  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 298)  
Gubin,A.N., Njoroge,J.M., Bouffard,G.G. and Miller,J.L.  
Gene expression in proliferating human erythroid cells  
JOURNAL  
Genomics 59 (2), 168-177 (1999)  
MEDLINE  
99339981  
COMMENT  
Contact: Jeffery L. Miller  
Laboratory of Chemical Biology  
National Institute of Diabetes and Digestive and Kidney Diseases  
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
20892, USA  
Tel: 301 402 2373  
Fax: 301 435 5148  
Email: jm7f@nih.gov  
DNA Sequencing and analyses by National Institutes of Health  
Intramural Sequencing Center (NISC).  
Plate: 44 row: d column: 12  
Seq primer: -21M13 forward primer (ABI).  
Location/Qualifiers  
1..298  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="ax44d12"

```

/clone_lib="Proliferating Human Erythroid Cells (LCB:ax
library)"
/sex="unknown"
/tissue_type="blood"
/cell_type="Erythroid Cells"
/cell_line="Primary Culture of Peripheral Blood
Mononuclear Cells"
/dev_stage="Progenitor; EPO responsive CD71++++"
/lab_host="SOLR"
/notes="Organ: blood; Vector: Lambda ZAP II; Site_1: EcoRI;
Site_2: EcoRI; 65,000 proliferating erythroid cells from
the buffy coat of a blood donation were obtained by flow
cyometric separation after a 5-day culture period in the
presence of erythropoietin. Total RNA was purified from
the sorted cell population using Trizol reagent. RNA (0.3
ug) was converted into double stranded cDNA using
Clontech's CapFinder cDNA Library Construction Kit
(Clontech) according to the manufacturer's protocol and
cloned into EcoRI digested Lambda Zap II vector
(Stratagene). The phage library was amplified once prior
to in vivo excision in SOLR cells. Individual colonies
were grown, and the cDNA inserts were sequenced in high
throughput (NIH intramural sequencing center
http://www.nisc.nih.gov/)."
BASE COUNT      83 a   74 c   77 g   64 t
ORIGIN
Query Match      2.6%; Score 44; DB 13; Length 298;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 141 GAGTCAAGACTAGCTGGCCACATGGTGAACCCCTACTCTCTA 184
|||||
Db 164 GAGTCAAGACTAGCTGGCCACATGGTGAACCCCTACTCTCTA 207
|||||

RESULT 41
BI031739
LOCUS      BI031739      341 bp      mRNA      linear      EST 14-JUN-2001
DEFINITION IL5-MT0266-100401-410-h08 MT0266 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BI031739
VERSION     BI031739.1 GI:14438369
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 341)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&st2=IL5-MT0266-
100401-410-h08&t3=2001-04-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 339.
Location/Qualifiers
1 . 345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0194"
/dev_stage="Adult"
/notes="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

```

```

1. .341
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MT0266"
/dev_stage="Adult"
/notes="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT      106 a   88 c   88 g   58 t   1 others
ORIGIN
Query Match      2.6%; Score 44; DB 13; Length 341;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 224 TAATCCCACTACTTAGGAGGCTGAGCGAGGAGAAATGCTTGAA 267
|||||
Db 196 TAATCCCACTACTTAGGAGGCTGAGCGAGGAGAAATGCTTGAA 239
|||||

RESULT 42
BG006754
LOCUS      BG006754      345 bp      mRNA      linear      EST 24-JAN-2001
DEFINITION RC3-GN0194-271100-023-C11 GN0194 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG006754
VERSION     BG006754.1 GI:12450255
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 345)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&st2=RC3-GN0194-
271100-023-C11&t3=2000-11-27&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 345.
Location/Qualifiers
1 . 345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="GN0194"
/dev_stage="Adult"
/notes="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
; Site_2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

```

FEATURES  
source

FEATURES



TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
BAC end sequences of Library RPCI-43  
Unpublished  
2 (bases 1 to 673)  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Totoki, Y., Watanabe, H. and Sakaki, Y.  
Direct Submission  
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC):  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimpbes@gsc.riken.go.jp, URL: http://ngp.gsc.riken.go.jp/  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the chimpanzee BAC library RPCI-43. This BAC  
end was generated during the R&D process and may have higher chance  
of clone tracking errors.  
PRIMERS

Sequencing: TJ

LIBRARY  
Vector : pBACe3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI  
Location/Qualifiers  
1. .673  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="RP43-024D09.TJ"  
/sex="male"  
/cell\_type="lymphocytes"  
/clone\_lib="RPCI-43 Chimpanzee Male BAC Library"  
199 a 118 c 154 g 201 t 1 others

FEATURES  
source

Query Match 2.6%; Score 44; DB 17; Length 673;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 224 TAATCCAGCTACTTAGGAGGCTGAGCGAGAGAAATTCGTTGAA 267  
|||||  
Db 634 TAATCCAGCTACTTAGGAGGCTGAGCGAGAGAAATTCGTTGAA 591  
|||||

RESULT 46  
AA878490/c 185 bp mRNA linear EST 25-MAR-1998  
LOCUS  
DEFINITION  
oel5b09.s1 NCI\_CGAP\_Ov2 Homo sapiens cDNA clone IMAGE:1385945  
similar to contains Alu repetitive element; contains element MER22  
repetitive element ;, mRNA sequence.  
ACCESSION  
AA878490  
VERSION  
AA878490.1 GI:2987455  
KEYWORDS  
EST.  
SOURCE  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1. (bases 1 to 185)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
Unpublished (1997)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue procurement: Christopher A. Moskaluk, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Sequencing by: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Seq primer: -40ml3 fwd. Et from Amersham  
High quality sequence stop: 113.  
Location/Qualifiers  
1. .185  
/organism="Homo sapiens"

FEATURES  
source

/db\_xref="taxon:9606"  
/clone="IMAGE:1385945"  
/clone\_lib="NCI\_CGAP\_Ov2"  
/sex="female"  
/tissue\_type="ovary"  
/lab\_host="DH108"  
/note="Vector: pAMP10; mRNA made from invasive ovarian  
tumor, cDNA made by oligo-dT priming. Non-directionally  
cloned. Size-selected on agarose gel, average insert size  
600 bp. Reference: Krizman et al. (1996) Cancer Research  
56:5380-5383."  
BASE COUNT 36 a 55 c 42 g 52 t  
ORIGIN

Query Match 2.6%; Score 43; DB 9; Length 185;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 CACACCTGTAATCCAGCTACTTAGGAGGCTGAGCGAGAGAA 258  
|||||  
Db 155 CACACCTGTAATCCAGCTACTTAGGAGGCTGAGCGAGAGAA 113  
|||||

RESULT 47  
BI054067 185 bp mRNA linear EST 15-JUN-2001  
LOCUS  
MR3-GN0349-300101-002-c05 GN0349 Homo sapiens cDNA, mRNA sequence.  
DEFINITION  
BI054067  
ACCESSION  
BI054067.1 GI:14461597  
VERSION  
EST.  
KEYWORDS  
human.  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1. (bases 1 to 185)  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Nagai, M.A., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
Goldman, G.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE  
20202663  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil

Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&st2=MR3-GN0349-300101-002-c05&t3=2001-01-30&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence stop: 185.  
Location/Qualifiers  
1. .185  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="GN0349"  
/dev\_stage="Adult"  
/note="Organ: placenta\_normal; Vector: puc18; Site\_1: SmaI  
; Site\_2: SmaI; A mini-library was made by cloning  
products derived from ORFEST PCR (U.S. Letters Patent  
application No. 196,716 - Ludwig Institute for Cancer  
Research) profiles into the pUC 18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."

FEATURES  
source

BASE COUNT 52 a 43 c 54 g 36 t



[illegible]

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-QV4-ST0233-251>)

199-041-a10&t3=1999-11-25&t4=1

Seq primer: puc 18 forward

High quality sequence stop: 430.

FEATURES

source

1. .430

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="ST0233"

/dev\_stage="Adult"

/note="Organ: stomach; Vector: puc18; Site\_1: SmaI;

Site\_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

BASE COUNT 124 a 74 c 122 g 110 t

ORIGIN

Query Match

Best Local Similarity 2.6%; Score 43; DB 10; Length 430;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 470 TGCCACTGCACCTCGGCGCTGGGCAACAGAGTGAGACCCCTGTCT 512

|||||

Db 426 TGCCACTGCACCTCGGCGCTGGGCAACAGAGTGAGACCCCTGTCT 384

Search completed: February 11, 2003, 11:06:25

Job time : 2615 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2003, 12:07:17 ; Search time 72 Seconds  
(without alignments)

7155.791 Million cell updates/sec

Title: US-09-942-310-2

Perfect score: 1680

Sequence: 1 gaattcaagaccagctgga.....catcttcctgctcgtgtgg 1680

Scoring table:

OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 20

Total number of hits satisfying chosen parameters: 6780

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61	3.6	180	6	Patent No. 5508199
2	61	3.6	1566	2	US-08-145-658D-13
3	61	3.6	1566	2	US-08-145-658D-22
4	61	3.6	1568	2	US-08-145-658D-20
5	61	3.6	1571	2	US-08-145-658D-21
6	42	2.5	84495	4	US-09-797-906-3
7	41	2.4	2115	1	US-08-395-800A-7
8	41	2.4	3715	4	US-09-085-199B-44
9	41	2.4	111282	4	US-09-754-250-3
10	38	2.3	11298	1	US-07-869-933-31
11	38	2.3	11298	4	US-08-201-879A-2
12	38	2.3	11298	4	US-09-103-663-31
13	38	2.3	112132	4	US-09-741-150-3
14	37	2.2	7210	2	US-08-257-963B-10
15	37	2.2	7210	4	US-08-367-841A-10
16	37	2.2	7210	5	PCT-US95-07201-10
17	37	2.2	14581	4	US-08-520-373D-4
18	37	2.2	22481	4	US-08-367-841A-43
19	37	2.2	22481	5	PCT-US95-07201-43
20	37	2.2	22484	4	US-09-875-223-2
21	37	2.2	80246	4	US-09-078-294-4
22	37	2.2	80595	4	US-09-078-294-3
23	37	2.2	162450	4	US-09-345-882-1
24	36	2.1	152331	3	US-09-128-155-16
25	36	2.1	168575	4	US-09-426-290-1
26	35	2.1	6235	4	US-09-305-384-5
27	35	2.1	6679	4	US-09-305-384-1

c 101	29	1.7	4326	2	US-08-852-807-12	Sequence 12, Appl	174	28	1.7	5761	1	US-08-463-335-2	Sequence 2, Appl
c 102	29	1.7	4460	4	US-09-103-875-4	Sequence 4, Appl	175	28	1.7	5781	2	US-08-464-023A-2	Sequence 2, Appl
c 103	29	1.7	7676	1	US-08-451-777A-7	Sequence 7, Appl	c 176	28	1.7	6078	3	US-08-816-617A-1	Sequence 1, Appl
c 104	29	1.7	7676	2	US-08-451-778A-7	Sequence 7, Appl	c 177	28	1.7	7042	4	US-09-092-508-1	Sequence 1, Appl
c 105	29	1.7	7676	2	US-08-998-208-7	Sequence 7, Appl	c 178	28	1.7	7042	4	US-09-435-115-1	Sequence 1, Appl
c 106	29	1.7	7676	5	PCT-US95-06743-7	Sequence 7, Appl	c 179	28	1.7	7042	4	US-09-098-310-1	Sequence 1, Appl
c 107	29	1.7	8758	4	US-09-799-345-3	Sequence 3, Appl	c 180	28	1.7	7042	4	US-09-690-364-21	Sequence 21, Appl
c 108	29	1.7	9301	4	US-09-449-218D-18	Sequence 18, Appl	c 181	28	1.7	7075	4	US-09-092-508-15	Sequence 15, Appl
c 109	29	1.7	13674	2	US-08-852-807-1	Sequence 1, Appl	c 182	28	1.7	7075	4	US-09-435-115-15	Sequence 15, Appl
c 110	29	1.7	17327	1	US-07-906-871-15	Sequence 15, Appl	c 183	28	1.7	8133	4	US-09-659-791A-10	Sequence 10, Appl
c 111	29	1.7	17327	1	US-07-906-871-15	Sequence 15, Appl	c 184	28	1.7	11811	4	US-09-078-294-7	Sequence 7, Appl
c 112	29	1.7	36159	4	US-09-749-368-3	Sequence 3, Appl	c 185	28	1.7	13875	2	US-08-734-344-1	Sequence 1, Appl
c 113	29	1.7	36651	4	US-09-738-894A-3	Sequence 3, Appl	c 186	28	1.7	16389	4	US-09-741-154-3	Sequence 3, Appl
c 114	29	1.7	36741	4	US-09-301-665-3	Sequence 3, Appl	c 187	28	1.7	17606	4	US-08-943-731-4	Sequence 4, Appl
c 115	29	1.7	50000	4	US-09-146-053-4	Sequence 4, Appl	c 188	28	1.7	18073	4	US-09-078-294-12	Sequence 12, Appl
c 116	29	1.7	62804	4	US-09-800-960-3	Sequence 3, Appl	c 189	28	1.7	18073	4	US-08-370-975B-6	Sequence 6, Appl
c 117	29	1.7	65042	4	US-09-784-316-3	Sequence 3, Appl	c 190	28	1.7	20303	1	US-09-641-638-651	Sequence 651, App
c 118	29	1.7	72928	3	US-09-009-913-1	Sequence 3, Appl	c 191	28	1.7	20674	4	US-09-262-773-210	Sequence 210, App
c 119	29	1.7	98844	4	US-09-791-211-10	Sequence 10, Appl	c 192	28	1.7	23071	4	US-08-370-975B-1	Sequence 1, Appl
c 120	29	1.7	99500	4	US-09-798-096-10	Sequence 10, Appl	c 193	28	1.7	26761	4	US-09-301-665-3	Sequence 3, Appl
c 121	29	1.7	152331	3	US-09-128-155-16	Sequence 16, Appl	c 194	28	1.7	26761	4	US-09-780-049-18	Sequence 18, Appl
c 122	29	1.7	169998	4	US-09-676-610B-24	Sequence 24, Appl	c 195	28	1.7	40000	4	US-09-810-347-3	Sequence 3, Appl
c 123	29	1.7	176373	3	US-09-128-155-17	Sequence 17, Appl	c 196	28	1.7	42571	4	US-09-146-053-6	Sequence 6, Appl
c 124	29	1.7	246240	2	US-08-724-394A-20	Sequence 20, Appl	c 197	28	1.7	45546	4	US-08-965-048-5	Sequence 5, Appl
c 125	29	1.7	246240	2	US-08-724-394A-21	Sequence 21, Appl	c 198	28	1.7	45716	4	US-08-965-048-6	Sequence 6, Appl
c 126	29	1.7	246240	2	US-08-724-394A-21	Sequence 21, Appl	c 199	28	1.7	45989	4	US-09-146-053-3	Sequence 3, Appl
c 127	29	1.7	246240	2	US-08-724-394A-22	Sequence 22, Appl	c 200	28	1.7	55827	4	US-09-813-133A-3	Sequence 3, Appl
c 128	29	1.7	246240	2	US-08-724-394A-22	Sequence 22, Appl	c 201	28	1.7	72604	4	US-09-268-992-7	Sequence 7, Appl
c 129	29	1.7	246240	2	US-08-724-394A-22	Sequence 22, Appl	c 202	28	1.7	72604	4	US-09-657-474-7	Sequence 7, Appl
c 130	28	1.7	280	2	US-08-849-701-7	Sequence 7, Appl	c 203	28	1.7	80246	4	US-09-078-294-4	Sequence 4, Appl
c 131	28	1.7	336	4	US-09-385-982-17	Sequence 17, Appl	c 204	28	1.7	80595	3	US-08-781-891-79	Sequence 79, Appl
c 132	28	1.7	489	4	US-09-370-838-109	Sequence 109, App	c 205	28	1.7	87350	3	US-09-078-294-3	Sequence 3, Appl
c 133	28	1.7	594	4	US-09-280-116-240	Sequence 240, App	c 206	28	1.7	87350	3	US-09-791-211-3	Sequence 3, Appl
c 134	28	1.7	652	4	US-09-328-111-717	Sequence 117, App	c 207	28	1.7	87343	4	US-09-791-211-3	Sequence 3, Appl
c 135	28	1.7	719	4	US-09-327-357-74	Sequence 74, Appl	c 208	28	1.7	98844	4	US-08-849-701-5	Sequence 5, Appl
c 136	28	1.7	837	1	US-08-832-883-56	Sequence 56, Appl	c 209	28	1.6	201	2	US-08-686-878A-15	Sequence 15, Appl
c 137	28	1.7	837	2	US-08-832-883-56	Sequence 56, Appl	c 210	27	1.6	218	1	US-09-222-575-88	Sequence 88, Appl
c 138	28	1.7	1001	4	US-09-641-638-98	Sequence 98, Appl	c 211	27	1.6	320	1	US-08-629-939-5	Sequence 5, Appl
c 139	28	1.7	1002	4	US-09-641-638-98	Sequence 98, Appl	c 212	27	1.6	320	1	US-08-759-873-5	Sequence 5, Appl
c 140	28	1.7	1253	2	US-08-591-629-1	Sequence 1, Appl	c 213	27	1.6	345	4	US-08-991-789A-214	Sequence 214, App
c 141	28	1.7	1260	4	US-08-943-731-169	Sequence 169, App	c 214	27	1.6	345	4	US-09-062-451-214	Sequence 214, App
c 142	28	1.7	1459	4	US-09-020-956-174	Sequence 174, App	c 215	27	1.6	345	4	US-09-598-326-214	Sequence 214, App
c 143	28	1.7	1459	4	US-09-030-607-174	Sequence 174, App	c 216	27	1.6	452	4	US-09-397-787-237	Sequence 237, App
c 144	28	1.7	1459	4	US-09-605-785-174	Sequence 174, App	c 217	27	1.6	550	4	US-08-998-416-148	Sequence 148, App
c 145	28	1.7	1459	4	US-09-439-313-174	Sequence 174, App	c 218	27	1.6	651	4	US-09-257-179-12	Sequence 12, Appl
c 146	28	1.7	1459	4	US-09-352-616A-174	Sequence 174, App	c 219	27	1.6	728	4	US-09-404-879A-16	Sequence 16, Appl
c 147	28	1.7	1459	4	US-09-232-149A-174	Sequence 174, App	c 220	27	1.6	821	4	US-09-342-681C-7	Sequence 7, Appl
c 148	28	1.7	1630	4	US-09-175-928-17	Sequence 17, App	c 221	27	1.6	856	4	US-09-288-143-47	Sequence 47, Appl
c 149	28	1.7	1808	1	US-08-351-149-4	Sequence 4, Appl	c 222	27	1.6	859	4	US-09-535-008-58	Sequence 58, Appl
c 150	28	1.7	1808	1	US-08-384-828-4	Sequence 4, Appl	c 223	27	1.6	940	4	US-09-069-023-9	Sequence 9, Appl
c 151	28	1.7	1808	3	US-08-895-474-4	Sequence 4, Appl	c 224	27	1.6	1001	4	US-09-641-638-121	Sequence 121, App
c 152	28	1.7	1856	1	US-08-157-171-3	Sequence 3, Appl	c 225	27	1.6	1001	4	US-09-641-638-448	Sequence 448, App
c 153	28	1.7	1856	4	US-09-050-159-128	Sequence 128, App	c 226	27	1.6	1037	4	US-09-257-179-16	Sequence 16, Appl
c 154	28	1.7	2238	1	US-08-742-011-1	Sequence 1, Appl	c 227	27	1.6	1050	3	US-08-755-587-21	Sequence 21, Appl
c 155	28	1.7	2245	6	5463025-3	Patent No. 5463025	c 228	27	1.6	1260	1	US-08-599-252-83	Sequence 83, Appl
c 156	28	1.7	2336	1	US-08-247-946A-1	Sequence 1, Appl	c 229	27	1.6	1260	1	US-08-436-074-56	Sequence 56, Appl
c 157	28	1.7	2336	5	PCT-US95-06420-1	Sequence 1, Appl	c 230	27	1.6	1260	5	PCT-US96-06352-83	Sequence 83, Appl
c 158	28	1.7	2448	2	US-08-687-080-111	Sequence 111, App	c 231	27	1.6	1260	5	PCT-US96-06352-83	Sequence 83, Appl
c 159	28	1.7	2713	2	US-08-916-901-6	Sequence 6, Appl	c 232	27	1.6	1607	4	US-09-091-087-9	Sequence 9, Appl
c 160	28	1.7	2713	4	US-09-154-602-6	Sequence 6, Appl	c 233	27	1.6	1762	4	US-09-443-184-35	Sequence 35, Appl
c 161	28	1.7	2713	3	US-09-038-832-1	Sequence 1, Appl	c 234	27	1.6	2343	2	US-09-031-392-1	Sequence 1, Appl
c 162	28	1.7	2900	3	US-09-262-773-209	Sequence 209, App	c 235	27	1.6	2343	3	US-09-299-549-1	Sequence 1, Appl
c 163	28	1.7	3607	2	US-08-629-001A-8	Sequence 8, Appl	c 236	27	1.6	2343	3	US-09-610-417-1	Sequence 1, Appl
c 164	28	1.7	3607	4	US-08-642-274D-8	Sequence 8, Appl	c 237	27	1.6	2415	3	US-09-019-689-1	Sequence 1, Appl
c 165	28	1.7	3607	4	US-08-952-127-8	Sequence 8, Appl	c 238	27	1.6	2415	3	US-09-019-689-1	Sequence 1, Appl
c 166	28	1.7	3607	4	US-08-952-014C-8	Sequence 8, Appl	c 239	27	1.6	2418	4	US-09-285-379-1	Sequence 1, Appl
c 167	28	1.7	3607	4	US-09-705-299-11	Sequence 11, Appl	c 240	27	1.6	2480	4	US-09-534-638-3	Sequence 3, Appl
c 168	28	1.7	3885	1	US-08-688-145-1	Sequence 1, Appl	c 241	27	1.6	2861	3	US-08-770-301A-12	Sequence 12, Appl
c 169	28	1.7	4285	4	US-09-040-774-1	Sequence 1, Appl	c 242	27	1.6	3011	1	US-07-821-716-1	Sequence 1, Appl
c 170	28	1.7	5543	2	US-08-687-080-101	Sequence 101, App	c 243	27	1.6	3366	1	US-08-469-802B-1	Sequence 1, Appl
c 171	28	1.7	5615	4	US-09-302-769-47	Sequence 47, Appl	c 244	27	1.6	3366	2	US-08-267-803B-1	Sequence 1, Appl
c 172	28	1.7	5761	1	US-07-749-001-2	Sequence 2, Appl	c 245	27	1.6	3742	1	US-08-694-915-5	Sequence 5, Appl
c 173	28	1.7	5761	1	US-08-154-198-2	Sequence 2, Appl	c 246	27	1.6	4668	4	US-09-045-301-1	Sequence 1, Appl

247	27	1.6	5141	1	US-08-337-690A-1	Sequence 1, Appli	320	26	1.5	1501	2	US-08-145-658D-24	Sequence 24, Appl
248	27	1.6	5141	4	US-09-048-887-1	Sequence 1, Appli	321	26	1.5	1624	4	US-08-430-225A-19	Sequence 19, Appl
c 249	27	1.6	6769	1	US-08-480-784-20	Sequence 20, Appl	c 322	26	1.5	1624	1	US-07-602-608-11	Sequence 11, Appl
c 250	27	1.6	6769	1	US-08-483-553-20	Sequence 20, Appl	c 323	26	1.5	2048	1	US-08-261-578-11	Sequence 11, Appl
c 251	27	1.6	6769	1	US-08-487-002-20	Sequence 20, Appl	c 324	26	1.5	2394	4	US-09-414-010-3	Sequence 3, Appli
c 252	27	1.6	6769	1	US-08-483-554B-20	Sequence 20, Appl	c 325	26	1.5	2509	2	US-09-014-969-1	Sequence 1, Appli
c 253	27	1.6	6769	1	US-08-488-011B-20	Sequence 20, Appl	c 326	26	1.5	2542	1	US-08-441-370-1	Sequence 4, Appli
c 254	27	1.6	6769	4	US-08-850-727-20	Sequence 20, Appl	c 327	26	1.5	2562	1	US-08-446-421-4	Sequence 4, Appli
c 255	27	1.6	6769	5	PCT-US95-10203-20	Sequence 20, Appl	c 328	26	1.5	2598	4	US-09-026-033-18	Sequence 18, Appl
c 256	27	1.6	6769	5	PCT-US95-10203-20	Sequence 20, Appl	c 329	26	1.5	3101	4	US-09-602-877A-97	Sequence 97, Appl
c 257	27	1.6	6769	5	PCT-US95-10202-20	Sequence 20, Appl	c 330	26	1.5	3441	4	US-09-026-033-17	Sequence 17, Appl
258	27	1.6	7452	3	US-08-592-500-1	Sequence 1, Appli	331	26	1.5	3630	3	US-08-434-000A-5	Sequence 5, Appli
259	27	1.6	7452	5	US-08-195-006-1	Sequence 1, Appli	332	26	1.5	3630	4	US-09-312-157-5	Sequence 5, Appli
260	27	1.6	7452	5	PCT-US94-07644A-1	Sequence 1, Appli	c 333	26	1.5	4129	2	US-08-370-319C-12	Sequence 12, Appl
261	27	1.6	13158	2	US-08-687-080-105	Sequence 105, App	c 334	26	1.5	4129	4	US-09-224-834-12	Sequence 12, Appl
c 262	27	1.6	13158	2	US-08-687-080-105	Sequence 105, App	c 335	26	1.5	4793	4	US-09-561-497-10	Sequence 10, Appl
c 263	27	1.6	14636	4	US-09-173-914-6	Sequence 6, Appli	c 336	26	1.5	6476	4	US-09-127-670-5	Sequence 5, Appli
c 264	27	1.6	14796	4	US-08-975-080-35	Sequence 35, Appl	337	26	1.5	6987	4	US-09-026-033-3	Sequence 3, Appli
c 265	27	1.6	14796	4	US-09-630-706-10	Sequence 10, Appl	338	26	1.5	6990	4	US-09-026-033-23	Sequence 23, Appl
c 266	27	1.6	14796	4	US-09-496-694B-3	Sequence 3, Appli	c 339	26	1.5	7210	2	US-08-257-963B-10	Sequence 10, Appl
c 267	27	1.6	16063	4	US-09-801-052-3	Sequence 3, Appli	c 340	26	1.5	7210	4	US-08-367-841A-10	Sequence 10, Appl
c 268	27	1.6	16389	4	US-09-741-154-3	Sequence 3, Appli	c 341	26	1.5	7210	5	PCT-US95-07201-10	Sequence 10, Appl
c 269	27	1.6	18073	4	US-09-078-294-12	Sequence 12, Appl	c 342	26	1.5	7620	1	US-07-767-135-1	Sequence 1, Appli
c 270	27	1.6	18443	4	US-09-078-294-6	Sequence 6, Appli	c 343	26	1.5	7620	1	US-07-841-652-1	Sequence 1, Appli
c 271	27	1.6	19736	2	US-09-740-035-3	Sequence 3, Appli	c 344	26	1.5	7720	4	US-09-318-448-5	Sequence 5, Appli
c 272	27	1.6	22846	2	US-08-469-461-3	Sequence 3, Appli	c 345	26	1.5	8056	3	US-09-136-605-14	Sequence 14, Appl
c 273	27	1.6	22846	3	US-07-890-609-3	Sequence 3, Appli	c 346	26	1.5	8082	1	US-08-306-691B-41	Sequence 41, Appl
274	27	1.6	36159	4	US-09-749-588-3	Sequence 3, Appli	c 347	26	1.5	8082	1	US-08-187-785-1	Sequence 1, Appli
c 275	27	1.6	38564	4	US-09-734-673-3	Sequence 3, Appli	c 348	26	1.5	8082	4	US-09-167-322-11	Sequence 11, Appl
c 276	27	1.6	42571	4	US-09-810-347-3	Sequence 3, Appli	c 349	26	1.5	8082	5	PCT-US93-06251-28	Sequence 28, Appl
c 277	27	1.6	43795	3	US-08-742-185-101	Sequence 101, App	c 350	26	1.5	8342	3	US-08-545-860D-63	Sequence 63, Appl
c 278	27	1.6	55827	4	US-09-813-133A-3	Sequence 3, Appli	351	26	1.5	8342	5	PCT-US94-04496-63	Sequence 6, Appli
c 279	27	1.6	62804	4	US-09-800-960-3	Sequence 3, Appli	352	26	1.5	8392	3	US-08-080-255-6	Sequence 6, Appli
c 280	27	1.6	65042	4	US-09-784-316-3	Sequence 3, Appli	353	26	1.5	8392	3	US-08-465-713-6	Sequence 6, Appli
c 281	26	1.5	56	2	US-08-776-944-9	Sequence 9, Appli	354	26	1.5	8392	5	PCT-US93-05857-6	Sequence 6, Appli
c 282	26	1.5	112	2	US-08-454-557C-27	Sequence 27, Appl	c 355	26	1.5	8517	4	US-08-827-208-1	Sequence 1, Appli
c 283	26	1.5	112	2	US-08-340-426D-27	Sequence 27, Appl	c 356	26	1.5	8517	4	US-09-500-358-1	Sequence 1, Appli
c 284	26	1.5	112	2	US-08-450-673C-27	Sequence 27, Appl	c 357	26	1.5	8517	4	US-09-498-809-1	Sequence 1, Appli
c 285	26	1.5	112	5	PCT-US95-17111A-27	Sequence 27, Appl	358	26	1.5	10684	3	US-08-618-100B-3	Sequence 3, Appli
c 286	26	1.5	322	4	US-09-385-982-216	Sequence 216, App	c 359	26	1.5	11464	3	US-08-884-324-13	Sequence 13, Appl
c 287	26	1.5	322	4	US-09-385-982-362	Sequence 362, App	c 360	26	1.5	12394	4	US-09-488-856A-10	Sequence 10, Appl
c 288	26	1.5	377	2	US-08-454-557C-37	Sequence 37, Appl	c 361	26	1.5	14581	4	US-08-520-373D-4	Sequence 4, Appli
289	26	1.5	377	2	US-08-340-426D-37	Sequence 37, Appl	c 362	26	1.5	16595	4	US-09-146-053-7	Sequence 7, Appli
290	26	1.5	377	2	US-08-450-673C-37	Sequence 37, Appl	c 363	26	1.5	19011	1	US-08-310-356-36	Sequence 36, Appl
291	26	1.5	377	5	PCT-US95-17111A-37	Sequence 37, Appl	c 364	26	1.5	19557	5	PCT-US92-06300-1	Sequence 1, Appli
292	26	1.5	382	4	US-09-438-906-23	Sequence 23, Appl	c 365	26	1.5	22481	5	US-08-367-841A-43	Sequence 43, Appl
c 293	26	1.5	403	4	US-09-385-982-29	Sequence 29, Appl	c 366	26	1.5	22481	5	PCT-US95-07201-43	Sequence 43, Appl
294	26	1.5	421	2	US-08-332-766A-25	Sequence 25, Appl	c 367	26	1.5	22484	4	US-09-875-223-2	Sequence 2, Appli
295	26	1.5	603	2	US-08-924-838-6	Sequence 6, Appli	c 368	26	1.5	28994	3	US-08-884-324-14	Sequence 14, Appl
c 296	26	1.5	619	4	US-09-152-060-17	Sequence 17, Appl	c 369	26	1.5	29629	4	US-09-729-995-3	Sequence 3, Appli
c 297	26	1.5	624	4	US-09-385-982-359	Sequence 359, App	c 370	26	1.5	40000	4	US-09-780-049-18	Sequence 18, Appl
298	26	1.5	685	4	US-09-227-357-100	Sequence 100, App	371	26	1.5	53526	3	US-08-658-136-2	Sequence 2, Appli
c 299	26	1.5	764	4	US-09-288-143-57	Sequence 57, Appl	372	26	1.5	53577	3	US-08-658-136-1	Sequence 1, Appli
c 300	26	1.5	774	3	US-08-755-587-20	Sequence 20, Appl	c 373	26	1.5	59065	4	US-09-813-817-3	Sequence 3, Appli
301	26	1.5	826	1	US-08-698-551-3	Sequence 3, Appli	c 374	26	1.5	59065	4	US-09-978-197-3	Sequence 3, Appli
302	26	1.5	826	2	US-08-602-228-3	Sequence 3, Appli	375	26	1.5	79628	3	US-09-009-913-1	Sequence 1, Appli
303	26	1.5	826	2	US-08-649-341A-3	Sequence 3, Appli	376	25	1.5	81001	4	US-09-750-580-1	Sequence 1, Appli
304	26	1.5	826	2	US-08-494-440B-3	Sequence 3, Appli	377	25	1.5	81001	4	US-08-778-794A-133	Sequence 133, App
305	26	1.5	826	2	US-08-533-901B-3	Sequence 3, Appli	378	25	1.5	80	1	US-07-920-281C-25	Sequence 25, Appl
306	26	1.5	826	2	US-08-839-032A-3	Sequence 3, Appli	379	25	1.5	80	4	US-08-466-277-25	Sequence 25, Appl
307	26	1.5	826	2	US-08-839-031A-3	Sequence 3, Appli	c 380	25	1.5	128	4	US-09-183-266A-10	Sequence 10, Appl
308	26	1.5	826	4	US-09-185-256C-3	Sequence 3, Appli	c 381	25	1.5	179	4	US-09-062-451-206	Sequence 206, App
309	26	1.5	826	5	PCT-US95-12724-3	Sequence 3, Appli	382	25	1.5	179	4	US-09-598-326-206	Sequence 206, App
310	26	1.5	1040	1	US-08-276-452A-91	Sequence 91, Appl	383	25	1.5	179	4	US-09-598-326-206	Sequence 206, App
311	26	1.5	1040	2	US-08-798-744-91	Sequence 91, Appl	384	25	1.5	236	1	US-08-599-252-99	Sequence 99, Appl
312	26	1.5	1078	4	US-09-452-239-41	Sequence 41, Appl	c 385	25	1.5	236	5	PCT-US96-06352-99	Sequence 99, Appl
313	26	1.5	1132	3	US-08-651-136C-21	Sequence 21, Appl	c 386	25	1.5	236	5	PCT-US96-06583-99	Sequence 99, Appl
314	26	1.5	1132	4	US-09-229-911A-21	Sequence 21, Appl	c 387	25	1.5	236	5	PCT-US96-06583-99	Sequence 99, Appl
315	26	1.5	1371	3	US-08-684-324-11	Sequence 11, Appl	c 388	25	1.5	256	2	US-09-385-982-187	Sequence 187, App
316	26	1.5	1480	2	US-08-454-557C-38	Sequence 38, Appl	c 389	25	1.5	265	2	US-08-849-701-1	Sequence 1, Appli
317	26	1.5	1480	2	US-08-340-426D-38	Sequence 38, Appl	c 390	25	1.5	294	2	US-08-481-658B-61	Sequence 61, Appl
318	26	1.5	1480	2	US-08-450-673C-38	Sequence 38, Appl	c 391	25	1.5	294	2	US-08-477-504A-61	Sequence 61, Appl
319	26	1.5	1480	5	PCT-US95-17111A-38	Sequence 38, Appl	c 392	25	1.5	294	2	US-08-486-756A-61	Sequence 61, Appl

C 393	25	1.5	294	2	US-08-485-862B-61	Sequence 61, Appl	466	25	1.5	1270	5	PCT-US94-11121-23	Sequence 23, Appl
C 394	25	1.5	294	3	US-08-787-739-61	Sequence 61, Appl	467	25	1.5	1287	4	US-09-564-805-217	Sequence 217, Appl
C 395	25	1.5	294	3	US-08-487-077A-61	Sequence 61, Appl	468	25	1.5	1320	2	US-08-695-412B-13	Sequence 13, Appl
C 396	25	1.5	294	3	US-08-485-863A-61	Sequence 61, Appl	469	25	1.5	1320	4	US-09-255-154D-13	Sequence 13, Appl
C 397	25	1.5	294	3	US-08-485-863A-61	Sequence 61, Appl	470	25	1.5	1355	4	US-09-370-838-31	Sequence 31, Appl
C 398	25	1.5	294	4	US-08-485-049D-61	Sequence 61, Appl	c 471	25	1.5	1355	4	US-08-481-658B-49	Sequence 49, Appl
C 399	25	1.5	294	4	US-09-178-115-61	Sequence 61, Appl	c 472	25	1.5	1401	2	US-08-477-504A-49	Sequence 49, Appl
C 400	25	1.5	294	4	US-09-177-776-61	Sequence 61, Appl	c 473	25	1.5	1401	2	US-08-486-756A-49	Sequence 49, Appl
C 401	25	1.5	328	4	US-09-605-785-418	Sequence 418, App	c 474	25	1.5	1401	2	US-08-485-862B-49	Sequence 49, Appl
C 402	25	1.5	328	4	US-09-352-616A-418	Sequence 418, App	c 475	25	1.5	1401	3	US-08-787-739-49	Sequence 49, Appl
C 403	25	1.5	374	4	US-09-385-982-135	Sequence 135, App	c 476	25	1.5	1401	3	US-08-487-077A-49	Sequence 49, Appl
C 404	25	1.5	425	2	US-08-967-101-102	Sequence 102, App	c 477	25	1.5	1401	3	US-08-485-863A-49	Sequence 49, Appl
C 405	25	1.5	425	2	US-08-592-541-102	Sequence 102, App	c 478	25	1.5	1401	3	US-08-485-049D-49	Sequence 49, Appl
C 406	25	1.5	425	3	US-09-124-698-102	Sequence 102, App	c 479	25	1.5	1401	4	US-09-178-115-49	Sequence 49, Appl
C 407	25	1.5	425	4	US-09-127-480-102	Sequence 102, App	c 480	25	1.5	1401	4	US-09-177-776-49	Sequence 49, Appl
C 408	25	1.5	425	4	US-09-496-841C-102	Sequence 102, App	481	25	1.5	1460	4	US-09-257-179-23	Sequence 23, Appl
C 409	25	1.5	425	4	US-09-124-523-102	Sequence 102, App	482	25	1.5	1493	1	US-08-340-820-24	Sequence 24, Appl
C 410	25	1.5	473	4	US-09-328-111-432	Sequence 432, App	483	25	1.5	1493	1	US-08-593-535-24	Sequence 24, Appl
C 411	25	1.5	500	2	US-08-475-844-18	Sequence 18, Appl	484	25	1.5	1525	1	US-08-186-833-3	Sequence 3, Appl
C 412	25	1.5	500	5	PCT-US95-08429-18	Sequence 18, Appl	485	25	1.5	1525	1	US-08-609-572-1	Sequence 1, Appl
C 413	25	1.5	506	4	US-09-149-476-252	Sequence 252, App	486	25	1.5	1525	4	US-08-841-751-1	Sequence 1, Appl
C 414	25	1.5	535	4	US-09-385-982-385	Sequence 385, App	487	25	1.5	1525	4	US-08-846-340-1	Sequence 1, Appl
C 415	25	1.5	580	2	US-08-809-763-2	Sequence 2, Appl	488	25	1.5	1525	4	US-08-846-344-1	Sequence 1, Appl
C 416	25	1.5	580	3	US-08-956-253-2	Sequence 2, Appl	489	25	1.5	1776	2	US-09-041-886-12	Sequence 12, Appl
C 417	25	1.5	608	4	US-09-385-982-523	Sequence 523, App	490	25	1.5	1776	4	US-09-370-253-5	Sequence 5, Appl
C 418	25	1.5	609	4	US-09-385-982-291	Sequence 291, App	491	25	1.5	1865	4	US-09-291-922-27	Sequence 27, Appl
C 419	25	1.5	613	4	US-09-385-982-144	Sequence 144, App	492	25	1.5	1872	4	US-09-469-242-3	Sequence 3, Appl
C 420	25	1.5	622	4	US-09-385-982-312	Sequence 312, App	493	25	1.5	1876	4	PCT-US93-05000-32	Sequence 32, Appl
C 421	25	1.5	630	4	US-09-342-681C-5	Sequence 5, Appl	c 494	25	1.5	1901	5	US-08-464-517-32	Sequence 32, Appl
C 422	25	1.5	632	4	US-09-328-111-798	Sequence 798, App	c 495	25	1.5	2022	2	US-08-246-361A-32	Sequence 32, Appl
C 423	25	1.5	655	4	US-09-288-143-20	Sequence 20, App	c 496	25	1.5	2022	2	US-08-463-772-32	Sequence 32, Appl
C 424	25	1.5	658	4	US-09-328-111-816	Sequence 816, App	c 497	25	1.5	2022	3	US-09-149-476-83	Sequence 83, Appl
C 425	25	1.5	669	4	US-09-328-111-782	Sequence 782, App	498	25	1.5	2024	4	US-09-500-063-1	Sequence 1, Appl
C 426	25	1.5	685	4	US-09-227-357-766	Sequence 66, App	499	25	1.5	2032	4	US-09-152-060-22	Sequence 22, Appl
C 427	25	1.5	688	6	5498694-3	Patent No. 5498694	500	25	1.5	2174	4	US-09-613-444-1	Sequence 1, Appl
C 428	25	1.5	704	4	US-08-896-164-49	Sequence 49, Appl	501	25	1.5	2214	4	US-08-943-731-57	Sequence 57, Appl
C 429	25	1.5	735	3	US-08-950-720A-5	Sequence 5, Appl	502	25	1.5	2232	4	US-09-212-609B-19	Sequence 19, Appl
C 430	25	1.5	807	2	US-08-531-927B-9	Sequence 9, Appl	503	25	1.5	2291	6	5281520-3	Patent No. 5281520
C 431	25	1.5	926	4	US-08-938-669A-4	Sequence 4, Appl	504	25	1.5	2310	1	US-08-471-570-9	Sequence 9, Appl
C 432	25	1.5	930	4	US-09-227-357-61	Sequence 61, Appl	505	25	1.5	2413	4	US-09-518-046-1	Sequence 1, Appl
C 433	25	1.5	930	4	US-09-227-357-146	Sequence 146, App	506	25	1.5	2416	4	US-09-261-416-1	Sequence 1, Appl
C 434	25	1.5	955	4	US-09-641-638-19	Sequence 19, Appl	507	25	1.5	2449	1	US-09-149-476-241	Sequence 241, App
C 435	25	1.5	955	4	US-09-641-638-20	Sequence 20, Appl	508	25	1.5	2461	1	US-08-832-883-3	Sequence 3, Appl
C 436	25	1.5	973	4	US-09-257-583-8	Sequence 8, Appl	509	25	1.5	2461	2	US-08-832-877-113	Sequence 113, App
C 437	25	1.5	1000	4	US-09-018-584A-30	Sequence 30, Appl	510	25	1.5	2544	2	US-09-518-046-3	Sequence 3, Appl
C 438	25	1.5	1000	4	US-09-018-584A-31	Sequence 31, Appl	511	25	1.5	2562	2	US-08-436-771-8	Sequence 8, Appl
C 439	25	1.5	1001	4	US-09-641-638-259	Sequence 259, App	512	25	1.5	2562	2	US-08-434-998-8	Sequence 8, Appl
C 440	25	1.5	1001	4	US-09-641-638-265	Sequence 265, App	513	25	1.5	2562	2	US-08-487-797-8	Sequence 8, Appl
C 441	25	1.5	1013	1	US-07-920-519-30	Sequence 30, Appl	514	25	1.5	2634	4	PCT-US95-02058-8	Sequence 8, Appl
C 442	25	1.5	1013	1	US-08-086-410-23	Sequence 23, Appl	515	25	1.5	2676	1	US-09-463-238-3	Sequence 3, Appl
C 443	25	1.5	1013	1	US-08-314-586-30	Sequence 30, Appl	516	25	1.5	2721	3	US-08-471-570-7	Sequence 7, Appl
C 444	25	1.5	1027	4	US-09-465-558-57	Sequence 57, Appl	517	25	1.5	2854	2	US-08-921-195-1	Sequence 1, Appl
C 445	25	1.5	1040	4	US-09-183-959-11	Sequence 11, Appl	518	25	1.5	3233	4	US-08-724-394A-17	Sequence 17, Appl
C 446	25	1.5	1064	1	US-08-378-588-15	Sequence 15, Appl	519	25	1.5	3401	2	US-08-671-975A-4	Sequence 4, Appl
C 447	25	1.5	1064	1	US-08-811-094-15	Sequence 15, Appl	520	25	1.5	3441	4	US-09-026-033-17	Sequence 17, Appl
C 448	25	1.5	1064	5	PCT-US94-11121-15	Sequence 15, Appl	521	25	1.5	3761	4	US-08-890-865A-2	Sequence 2, Appl
C 449	25	1.5	1116	3	US-08-672-569-1	Sequence 1, Appl	522	25	1.5	4004	4	US-09-293-505-8	Sequence 8, Appl
C 450	25	1.5	1116	3	US-08-916-443A-16	Sequence 16, Appl	523	25	1.5	4192	4	US-09-122-126B-1	Sequence 1, Appl
C 451	25	1.5	1174	2	US-08-481-658B-39	Sequence 39, Appl	524	25	1.5	4203	2	US-08-866-757-1	Sequence 1, Appl
C 452	25	1.5	1174	2	US-08-477-504A-39	Sequence 39, Appl	525	25	1.5	4203	4	US-09-153-593-1	Sequence 1, Appl
C 453	25	1.5	1174	2	US-08-486-756A-39	Sequence 39, Appl	526	25	1.5	4316	1	US-08-317-450B-14	Sequence 14, Appl
C 454	25	1.5	1174	2	US-08-485-862B-39	Sequence 39, Appl	527	25	1.5	4316	3	US-08-800-593-14	Sequence 14, Appl
C 455	25	1.5	1174	3	US-08-787-739-39	Sequence 39, Appl	528	25	1.5	4335	3	US-09-058-489-19	Sequence 19, Appl
C 456	25	1.5	1174	3	US-08-487-077A-39	Sequence 39, Appl	529	25	1.5	4671	4	US-08-462-437-27	Sequence 27, Appl
C 457	25	1.5	1174	3	US-08-485-863A-39	Sequence 39, Appl	530	25	1.5	5605	4	US-09-268-140-6	Sequence 6, Appl
C 458	25	1.5	1174	4	US-08-485-049D-39	Sequence 39, Appl	531	25	1.5	5605	4	US-09-305-639-4	Sequence 4, Appl
C 459	25	1.5	1174	4	US-09-178-115-39	Sequence 39, Appl	532	25	1.5	7622	4	US-09-305-639-1	Sequence 1, Appl
C 460	25	1.5	1174	4	US-09-177-776-39	Sequence 39, Appl	533	25	1.5	9844	4	US-08-462-437-30	Sequence 30, Appl
C 461	25	1.5	1208	1	US-07-949-516A-3	Sequence 3, Appl	534	25	1.5	10380	4	US-09-077-354B-3	Sequence 3, Appl
C 462	25	1.5	1208	2	US-08-814-459-3	Sequence 3, Appl	535	25	1.5	10642	4	US-09-934-551-3	Sequence 1, Appl
C 463	25	1.5	1208	3	US-09-122-525-3	Sequence 23, Appl	536	25	1.5	10825	3	US-08-652-265-1	Sequence 1, Appl
C 464	25	1.5	1270	1	US-08-378-588-23	Sequence 23, Appl	537	25	1.5				
C 465	25	1.5	1270	2	US-08-811-094-23	Sequence 23, Appl	538	25	1.5				

c 539	25	1.5	10825	3	US-08-652-265-3	Sequence 3, Appl	Sequence 3, Appl	c 612	24	1.4	450	2	US-08-592-541-7	Sequence 7, Appl
c 540	25	1.5	10825	3	US-08-652-265-5	Sequence 5, Appl	Sequence 5, Appl	c 613	24	1.4	450	3	US-09-124-698-7	Sequence 7, Appl
c 541	25	1.5	10825	3	US-08-652-265-7	Sequence 7, Appl	Sequence 7, Appl	c 614	24	1.4	450	4	US-09-127-480-7	Sequence 7, Appl
c 542	25	1.5	10825	3	US-08-834-497A-1	Sequence 1, Appl	Sequence 1, Appl	c 615	24	1.4	450	4	US-08-496-841C-7	Sequence 7, Appl
c 543	25	1.5	10825	3	US-08-834-497A-3	Sequence 3, Appl	Sequence 3, Appl	c 616	24	1.4	450	4	US-09-124-523-7	Sequence 7, Appl
c 544	25	1.5	10825	3	US-08-834-497A-5	Sequence 5, Appl	Sequence 5, Appl	c 617	24	1.4	470	4	US-09-020-956-102	Sequence 102, App
c 545	25	1.5	10825	3	US-08-834-497A-7	Sequence 7, Appl	Sequence 7, Appl	c 618	24	1.4	470	4	US-09-030-607-102	Sequence 102, App
c 546	25	1.5	10825	4	US-09-503-444A-1	Sequence 1, Appl	Sequence 1, Appl	c 619	24	1.4	470	4	US-09-605-785-102	Sequence 102, App
c 547	25	1.5	10825	4	US-09-503-444A-3	Sequence 3, Appl	Sequence 3, Appl	c 620	24	1.4	470	4	US-09-439-313-102	Sequence 102, App
c 548	25	1.5	10825	4	US-09-503-444A-5	Sequence 5, Appl	Sequence 5, Appl	c 621	24	1.4	470	4	US-09-352-616A-102	Sequence 102, App
c 549	25	1.5	10825	4	US-09-503-444A-7	Sequence 7, Appl	Sequence 7, Appl	c 622	24	1.4	470	4	US-09-232-149A-102	Sequence 102, App
c 550	25	1.5	10898	2	US-08-481-658B-5	Sequence 5, Appl	Sequence 5, Appl	c 623	24	1.4	492	4	US-09-280-116-208	Sequence 208, App
c 551	25	1.5	10898	2	US-08-477-504A-5	Sequence 5, Appl	Sequence 5, Appl	c 624	24	1.4	509	4	US-09-030-607-202	Sequence 202, App
c 552	25	1.5	10898	2	US-08-486-756A-5	Sequence 5, Appl	Sequence 5, Appl	c 625	24	1.4	509	4	US-09-605-785-202	Sequence 202, App
c 553	25	1.5	10898	2	US-08-485-862B-5	Sequence 5, Appl	Sequence 5, Appl	c 626	24	1.4	509	4	US-09-439-313-202	Sequence 202, App
c 554	25	1.5	10898	3	US-08-787-739-5	Sequence 5, Appl	Sequence 5, Appl	c 627	24	1.4	509	4	US-09-352-616A-202	Sequence 202, App
c 555	25	1.5	10898	3	US-08-487-077A-5	Sequence 5, Appl	Sequence 5, Appl	c 628	24	1.4	509	4	US-09-232-149A-202	Sequence 202, App
c 556	25	1.5	10898	3	US-08-485-863A-5	Sequence 5, Appl	Sequence 5, Appl	c 629	24	1.4	541	4	US-09-288-143-50	Sequence 50, Appl
c 557	25	1.5	10898	4	US-08-485-049D-5	Sequence 5, Appl	Sequence 5, Appl	c 630	24	1.4	566	4	US-09-328-111-466	Sequence 466, App
c 558	25	1.5	10898	4	US-09-178-115-5	Sequence 5, Appl	Sequence 5, Appl	c 631	24	1.4	569	4	US-09-227-357-89	Sequence 89, Appl
c 559	25	1.5	10898	4	US-09-177-776-5	Sequence 5, Appl	Sequence 5, Appl	c 632	24	1.4	571	1	US-08-322-742-14	Sequence 14, Appl
c 560	25	1.5	11517	1	US-07-920-281C-1	Sequence 1, Appl	Sequence 1, Appl	c 633	24	1.4	578	4	US-09-328-111-757	Sequence 757, App
c 561	25	1.5	11517	4	US-08-466-277-1	Sequence 1, Appl	Sequence 1, Appl	c 634	24	1.4	581	4	US-09-385-982-12	Sequence 12, Appl
c 562	25	1.5	12146	4	US-09-277-457-27	Sequence 27, Appl	Sequence 27, Appl	c 635	24	1.4	591	4	US-09-385-982-406	Sequence 406, App
c 563	25	1.5	12594	4	US-09-488-856A-10	Sequence 10, Appl	Sequence 10, Appl	c 636	24	1.4	606	4	US-09-040-984-55	Sequence 55, Appl
c 564	25	1.5	14753	4	US-09-821-736-3	Sequence 3, Appl	Sequence 3, Appl	c 637	24	1.4	606	4	US-09-123-912-55	Sequence 55, Appl
c 565	25	1.5	17041	1	US-08-076-011-1	Sequence 1, Appl	Sequence 1, Appl	c 638	24	1.4	606	4	US-09-643-597-55	Sequence 55, Appl
c 566	25	1.5	18596	4	US-09-318-448-11	Sequence 11, Appl	Sequence 11, Appl	c 639	24	1.4	607	4	US-09-385-982-288	Sequence 288, App
c 567	25	1.5	18609	4	US-08-943-731-1	Sequence 1, Appl	Sequence 1, Appl	c 640	24	1.4	608	4	US-09-385-982-183	Sequence 183, App
c 568	25	1.5	26664	4	US-09-564-805-28	Sequence 28, Appl	Sequence 28, Appl	c 641	24	1.4	611	4	US-09-385-982-178	Sequence 178, App
c 569	25	1.5	28001	4	US-09-819-993-3	Sequence 3, Appl	Sequence 3, Appl	c 642	24	1.4	611	4	US-09-385-982-393	Sequence 393, App
c 570	25	1.5	35060	3	US-09-814-095-7	Sequence 7, Appl	Sequence 7, Appl	c 643	24	1.4	622	4	US-09-385-982-184	Sequence 184, App
c 571	25	1.5	38564	4	US-09-734-673-3	Sequence 3, Appl	Sequence 3, Appl	c 644	24	1.4	629	4	US-09-385-982-4	Sequence 4, Appl
c 572	25	1.5	45716	4	US-08-965-048-5	Sequence 5, Appl	Sequence 5, Appl	c 645	24	1.4	629	4	US-09-385-982-305	Sequence 305, App
c 573	25	1.5	45989	4	US-08-965-048-6	Sequence 6, Appl	Sequence 6, Appl	c 646	24	1.4	631	4	US-09-385-982-354	Sequence 354, App
c 574	25	1.5	50000	4	US-09-146-053-4	Sequence 4, Appl	Sequence 4, Appl	c 647	24	1.4	632	4	US-09-385-982-499	Sequence 499, App
c 575	25	1.5	56516	2	US-08-996-306-1	Sequence 1, Appl	Sequence 1, Appl	c 648	24	1.4	635	1	US-08-455-633A-35	Sequence 35, Appl
c 576	25	1.5	56516	4	US-09-338-907-1	Sequence 1, Appl	Sequence 1, Appl	c 649	24	1.4	635	1	US-08-416-336-5	Sequence 35, Appl
c 577	25	1.5	56516	4	US-09-218-207-1	Sequence 1, Appl	Sequence 1, Appl	c 650	24	1.4	635	2	US-08-456-460C-35	Sequence 35, Appl
c 578	25	1.5	56520	4	US-09-338-907-179	Sequence 179, App	Sequence 179, App	c 651	24	1.4	635	5	PCT-US94-05354-35	Sequence 35, Appl
c 579	25	1.5	56520	4	US-09-218-207-179	Sequence 179, App	Sequence 179, App	c 652	24	1.4	645	4	US-09-328-111-106	Sequence 106, App
c 580	25	1.5	72604	4	US-09-268-992-7	Sequence 7, Appl	Sequence 7, Appl	c 653	24	1.4	653	4	US-09-373-750-1	Sequence 1, Appl
c 581	25	1.5	72604	4	US-09-657-474-7	Sequence 7, Appl	Sequence 7, Appl	c 654	24	1.4	657	4	US-09-385-982-91	Sequence 91, Appl
c 582	25	1.5	169998	4	US-09-676-610B-24	Sequence 24, Appl	Sequence 24, Appl	c 655	24	1.4	658	4	US-09-385-982-327	Sequence 327, App
c 583	24	1.4	38	4	US-09-325-554-7	Sequence 7, Appl	Sequence 7, Appl	c 656	24	1.4	660	1	US-08-555-678-41	Sequence 41, Appl
c 584	24	1.4	40	4	US-09-306-290-2	Sequence 2, Appl	Sequence 2, Appl	c 657	24	1.4	661	2	US-08-529-878B-37	Sequence 37, Appl
c 585	24	1.4	40	4	US-09-306-290-9	Sequence 9, Appl	Sequence 9, Appl	c 658	24	1.4	675	4	US-09-605-785-822	Sequence 822, App
c 586	24	1.4	55	2	US-08-771-624B-8	Sequence 8, Appl	Sequence 8, Appl	c 659	24	1.4	688	1	US-08-599-252-94	Sequence 94, Appl
c 587	24	1.4	75	2	US-08-776-944-13	Sequence 13, Appl	Sequence 13, Appl	c 660	24	1.4	688	5	PCT-US96-06352-94	Sequence 94, Appl
c 588	24	1.4	130	6	5198345-15	Patent No. 5198345	Patent No. 5198345	c 661	24	1.4	688	5	PCT-US96-06583-94	Sequence 94, Appl
c 589	24	1.4	130	6	5198345-15	Patent No. 5198345	Patent No. 5198345	c 662	24	1.4	689	4	US-09-105-542A-14	Sequence 14, Appl
c 590	24	1.4	194	3	US-08-951-200A-7	Sequence 7, Appl	Sequence 7, Appl	c 663	24	1.4	690	4	US-09-328-111-74	Sequence 74, Appl
c 591	24	1.4	218	4	US-09-480-921B-18	Sequence 18, Appl	Sequence 18, Appl	c 664	24	1.4	704	4	US-09-122-400B-8	Sequence 8, Appl
c 592	24	1.4	263	4	US-09-091-097-26	Sequence 26, Appl	Sequence 26, Appl	c 665	24	1.4	706	4	US-09-191-136-14	Sequence 14, Appl
c 593	24	1.4	282	1	US-08-133-629-8	Sequence 8, Appl	Sequence 8, Appl	c 666	24	1.4	712	4	US-09-149-476-318	Sequence 318, App
c 594	24	1.4	291	4	US-09-605-785-823	Sequence 823, App	Sequence 823, App	c 667	24	1.4	713	4	US-08-943-607-23	Sequence 23, Appl
c 595	24	1.4	294	1	US-08-446-660-18	Sequence 18, Appl	Sequence 18, Appl	c 668	24	1.4	713	4	US-08-943-607-24	Sequence 24, Appl
c 596	24	1.4	294	4	US-08-974-302-18	Sequence 18, Appl	Sequence 18, Appl	c 669	24	1.4	713	4	US-08-943-607-25	Sequence 25, Appl
c 597	24	1.4	296	4	US-09-385-982-19	Sequence 19, Appl	Sequence 19, Appl	c 670	24	1.4	713	4	US-08-943-607-26	Sequence 26, Appl
c 598	24	1.4	301	4	US-09-605-785-299	Sequence 299, App	Sequence 299, App	c 671	24	1.4	731	1	US-08-451-405A-2	Sequence 2, Appl
c 599	24	1.4	301	4	US-09-439-313-299	Sequence 299, App	Sequence 299, App	c 672	24	1.4	732	4	US-09-328-111-697	Sequence 697, App
c 600	24	1.4	301	4	US-09-352-616A-299	Sequence 299, App	Sequence 299, App	c 673	24	1.4	737	2	US-08-257-963B-41	Sequence 41, Appl
c 601	24	1.4	301	4	US-09-232-149A-299	Sequence 299, App	Sequence 299, App	c 674	24	1.4	737	4	US-08-367-841A-41	Sequence 41, Appl
c 602	24	1.4	327	4	US-09-385-982-544	Sequence 544, App	Sequence 544, App	c 675	24	1.4	737	5	PCT-US95-07201-41	Sequence 41, Appl
c 603	24	1.4	330	4	US-09-078-294-24	Sequence 24, Appl	Sequence 24, Appl	c 676	24	1.4	773	4	US-09-149-476-20	Sequence 20, Appl
c 604	24	1.4	336	4	US-09-385-982-508	Sequence 508, App	Sequence 508, App	c 677	24	1.4	774	3	US-08-765-340-1	Sequence 1, Appl
c 605	24	1.4	342	4	US-09-385-982-342	Sequence 342, App	Sequence 342, App	c 678	24	1.4	787	1	US-08-236-427-12	Sequence 12, Appl
c 606	24	1.4	352	4	US-09-385-982-529	Sequence 529, App	Sequence 529, App	c 679	24	1.4	798	4	US-09-288-143-21	Sequence 21, Appl
c 607	24	1.4	374	2	US-08-370-156-24	Sequence 24, Appl	Sequence 24, Appl	c 680	24	1.4	817	1	US-08-672-569-2	Sequence 2, Appl
c 608	24	1.4	380	4	US-09-385-982-457	Sequence 457, App	Sequence 457, App	c 681	24	1.4	826	4	US-09-288-143-45	Sequence 45, Appl
c 609	24	1.4	390	4	US-09-385-982-232	Sequence 232, App	Sequence 232, App	c 682	24	1.4	834	2	US-08-967-101-113	Sequence 113, App
c 610	24	1.4	437	4	US-09-091-725-46	Sequence 46, Appl	Sequence 46, Appl	c 683	24	1.4	834	3	US-08-592-541-113	Sequence 113, App
c 611	24	1.4	450	2	US-08-967-101-7	Sequence 7, Appl	Sequence 7, Appl	c 684	24	1.4	834	3	US-09-124-698-113	Sequence 113, App

685	24	1.4	834	4	US-09-127-480-113	Sequence 113, App	758	24	1.4	1781	1	US-08-481-130-4	Sequence 4, Appli
686	24	1.4	834	4	US-08-496-841C-113	Sequence 113, App	759	24	1.4	1781	1	US-08-656-984A-4	Sequence 4, Appli
687	24	1.4	834	4	US-09-124-523-113	Sequence 113, App	760	24	1.4	1781	1	US-08-482-882-2	Sequence 2, Appli
688	24	1.4	840	4	US-09-376-728-1	Sequence 1, Appli	761	24	1.4	1781	1	US-08-485-604-4	Sequence 4, Appli
689	24	1.4	863	4	US-09-697-367-3	Sequence 3, Appli	762	24	1.4	1781	1	US-08-483-389-2	Sequence 2, Appli
690	24	1.4	867	1	US-08-033-857A-9	Sequence 9, Appli	763	24	1.4	1781	2	US-08-487-113D-2	Sequence 2, Appli
691	24	1.4	867	1	US-08-374-983A-9	Sequence 9, Appli	764	24	1.4	1781	2	US-08-487-595-4	Sequence 2, Appli
692	24	1.4	867	1	US-08-374-983A-13	Sequence 13, Appli	765	24	1.4	1781	2	US-08-473-503-2	Sequence 2, Appli
693	24	1.4	868	3	US-08-889-502-20	Sequence 20, Appli	766	24	1.4	1781	2	US-08-483-932-2	Sequence 2, Appli
694	24	1.4	896	4	US-08-943-731-31	Sequence 31, Appli	767	24	1.4	1781	2	US-08-720-420A-2	Sequence 2, Appli
695	24	1.4	903	4	US-09-589-287B-18	Sequence 18, Appli	768	24	1.4	1781	3	US-08-714-017-2	Sequence 2, Appli
696	24	1.4	907	4	US-09-149-476-195	Sequence 195, App	769	24	1.4	1781	3	US-08-863-790-2	Sequence 2, Appli
697	24	1.4	911	4	US-09-364-230-25	Sequence 25, Appli	770	24	1.4	1781	3	US-08-475-680-2	Sequence 2, Appli
698	24	1.4	919	2	US-08-924-759-19	Sequence 19, Appli	771	24	1.4	1781	3	US-08-296-749-2	Sequence 2, Appli
699	24	1.4	919	3	US-09-248-335-19	Sequence 19, App	772	24	1.4	1797	4	US-09-157-603-2	Sequence 2, Appli
700	24	1.4	952	4	US-09-149-476-117	Sequence 117, App	773	24	1.4	1797	4	US-09-587-436-2	Sequence 2, Appli
701	24	1.4	971	4	US-08-535-008-39	Sequence 39, App	774	24	1.4	1797	4	US-08-927-165A-2	Sequence 2, Appli
702	24	1.4	1001	4	US-09-641-638-115	Sequence 115, App	775	24	1.4	1798	4	US-09-446-402A-16	Sequence 16, Appli
703	24	1.4	1001	4	US-09-641-638-464	Sequence 464, App	776	24	1.4	1799	4	US-09-329-633A-1	Sequence 1, Appli
c 704	24	1.4	1001	4	US-09-641-638-510	Sequence 510, App	777	24	1.4	1799	4	US-09-079-029-2	Sequence 2, Appli
705	24	1.4	1001	4	US-09-641-638-526	Sequence 526, App	778	24	1.4	1814	4	US-09-347-801-9	Sequence 9, Appli
c 706	24	1.4	1019	4	US-09-177-650-128	Sequence 128, App	779	24	1.4	1831	2	US-08-101-435-3	Sequence 3, Appli
707	24	1.4	1033	4	US-09-173-300-46	Sequence 46, Appli	780	24	1.4	1855	3	US-08-928-613-1	Sequence 1, Appli
c 708	24	1.4	1096	4	US-09-000-127-2	Sequence 2, Appli	781	24	1.4	1855	3	US-09-193-524-1	Sequence 1, Appli
709	24	1.4	1106	3	US-08-755-587-18	Sequence 18, Appli	782	24	1.4	1857	4	US-09-377-557-9	Sequence 9, Appli
c 710	24	1.4	1116	4	US-09-177-325-1	Sequence 1, Appli	783	24	1.4	1868	4	US-09-739-455-1	Sequence 1, Appli
c 711	24	1.4	1116	4	US-09-411-812A-1	Sequence 1, Appli	784	24	1.4	1878	4	US-09-732-025-1	Sequence 1, Appli
c 712	24	1.4	1116	4	US-09-590-113-1	Sequence 1, Appli	785	24	1.4	1882	3	US-08-501-572-4	Sequence 1, Appli
713	24	1.4	1128	3	US-09-136-628-1	Sequence 1, Appli	786	24	1.4	1882	3	US-09-040-444-4	Sequence 4, Appli
714	24	1.4	1215	2	US-08-370-156-26	Sequence 26, Appli	787	24	1.4	1894	2	US-08-935-450-7	Sequence 7, Appli
715	24	1.4	1278	4	US-09-008-697A-11	Sequence 11, App	788	24	1.4	1906	1	US-08-207-904-18	Sequence 18, Appli
716	24	1.4	1288	4	US-09-724-864-16	Sequence 16, Appli	789	24	1.4	1928	4	US-09-008-481A-9	Sequence 9, Appli
717	24	1.4	1305	4	US-09-242-859A-3	Sequence 3, Appli	790	24	1.4	1928	3	US-09-195-666A-15	Sequence 15, Appli
718	24	1.4	1305	4	US-09-242-859A-7	Sequence 7, Appli	791	24	1.4	1928	4	US-09-309-592-9	Sequence 9, Appli
719	24	1.4	1333	3	US-09-372-422A-9	Sequence 9, Appli	792	24	1.4	1928	4	US-09-635-705-15	Sequence 15, Appli
c 720	24	1.4	1334	3	US-08-884-324-8	Sequence 8, Appli	793	24	1.4	1928	4	US-09-634-858A-15	Sequence 15, Appli
721	24	1.4	1338	4	US-09-336-536-1	Sequence 1, Appli	794	24	1.4	1928	4	US-08-592-126-93	Sequence 93, Appli
722	24	1.4	1348	4	US-09-152-060-14	Sequence 14, Appli	795	24	1.4	1950	1	US-07-937-609-25	Sequence 25, Appli
723	24	1.4	1363	1	US-08-776-088-21	Sequence 21, Appli	796	24	1.4	1986	4	US-08-029-170-25	Sequence 25, Appli
724	24	1.4	1363	5	PCT-US95-09145A-21	Sequence 21, Appli	797	24	1.4	2015	1	US-08-029-170-25	Sequence 25, Appli
725	24	1.4	1386	2	US-08-897-340-1	Sequence 1, Appli	798	24	1.4	2015	4	US-07-937-609-25	Sequence 25, Appli
726	24	1.4	1386	3	US-09-252-329-1	Sequence 1, Appli	799	24	1.4	2030	3	US-08-706-216-3	Sequence 3, Appli
727	24	1.4	1391	1	US-08-261-662-1	Sequence 1, Appli	800	24	1.4	2114	4	US-09-130-491-7	Sequence 7, Appli
728	24	1.4	1391	5	PCT-US95-07752-1	Sequence 1, Appli	801	24	1.4	2125	3	US-09-109-204-5	Sequence 5, Appli
729	24	1.4	1396	1	US-08-123-161A-11	Sequence 11, Appli	802	24	1.4	2125	3	US-09-303-689-6	Sequence 6, Appli
730	24	1.4	1396	1	US-08-483-278-11	Sequence 11, Appli	803	24	1.4	2127	1	US-08-832-877-54	Sequence 54, Appli
731	24	1.4	1404	2	US-08-487-113D-119	Sequence 119, App	804	24	1.4	2158	1	US-08-832-877-54	Sequence 54, Appli
732	24	1.4	1404	2	US-08-720-420A-119	Sequence 119, App	805	24	1.4	2158	1	US-08-698-551-1	Sequence 1, Appli
733	24	1.4	1428	4	US-09-118-442-5	Sequence 5, Appli	806	24	1.4	2158	2	US-08-603-228-1	Sequence 1, Appli
734	24	1.4	1428	4	US-09-677-064-5	Sequence 5, Appli	807	24	1.4	2158	2	US-08-649-341A-1	Sequence 1, Appli
735	24	1.4	1469	5	PCT-US92-08258-1	Sequence 1, Appli	808	24	1.4	2158	2	US-08-494-440B-1	Sequence 1, Appli
736	24	1.4	1512	2	US-08-909-965C-8	Sequence 8, Appli	809	24	1.4	2158	2	US-08-533-901B-1	Sequence 1, Appli
737	24	1.4	1525	4	US-09-157-603-4	Sequence 4, Appli	810	24	1.4	2158	2	US-08-839-032A-1	Sequence 1, Appli
738	24	1.4	1525	4	US-09-587-436-4	Sequence 4, Appli	811	24	1.4	2158	2	US-08-839-031A-1	Sequence 1, Appli
739	24	1.4	1525	4	US-08-927-165A-4	Sequence 4, Appli	812	24	1.4	2158	5	PCT-US95-12724-1	Sequence 1, Appli
c 740	24	1.4	1541	6	5183884-1	Patent No. 5183884	813	24	1.4	2158	5	PCT-US95-12724-1	Sequence 1, Appli
c 741	24	1.4	1542	1	US-07-978-895-1	Sequence 1, Appli	814	24	1.4	2171	4	US-08-811-461-1	Sequence 1, Appli
c 742	24	1.4	1542	1	US-08-473-119-1	Sequence 1, Appli	815	24	1.4	2226	2	US-08-031-538-1	Sequence 1, Appli
c 743	24	1.4	1542	2	US-08-475-352-1	Sequence 1, Appli	816	24	1.4	2226	2	US-09-363-708-3	Sequence 3, Appli
744	24	1.4	1575	4	US-08-639-294-1	Sequence 1, Appli	317	24	1.4	2254	4	US-09-138-277C-2	Sequence 2, Appli
745	24	1.4	1609	4	US-09-342-647-9	Sequence 9, Appli	318	24	1.4	2254	4	US-08-318-826A-5	Sequence 5, Appli
746	24	1.4	1624	2	US-08-888-429A-17	Sequence 17, Appli	319	24	1.4	2256	2	US-08-370-156-1	Sequence 1, Appli
c 747	24	1.4	1624	2	US-08-852-807-10	Sequence 10, Appli	320	24	1.4	2256	3	US-08-814-095-1	Sequence 1, Appli
748	24	1.4	1631	1	US-08-583-318-4	Sequence 4, Appli	321	24	1.4	2256	3	US-08-788-750-1	Sequence 1, Appli
749	24	1.4	1643	3	US-08-888-429A-27	Sequence 27, Appli	822	24	1.4	2260	2	US-09-851-896-11	Sequence 11, Appli
750	24	1.4	1693	2	US-08-487-113D-118	Sequence 118, App	823	24	1.4	2328	4	US-08-811-481-34	Sequence 34, Appli
751	24	1.4	1693	2	US-08-720-420A-118	Sequence 118, App	824	24	1.4	2387	4	US-09-375-318-38	Sequence 38, Appli
752	24	1.4	1705	4	US-09-364-230-15	Sequence 15, App	825	24	1.4	2426	4	US-09-605-785-470	Sequence 470, App
753	24	1.4	1733	3	US-09-073-569-1	Sequence 1, Appli	826	24	1.4	2426	4	US-09-439-313-470	Sequence 470, App
754	24	1.4	1781	1	US-08-314-615-2	Sequence 2, Appli	827	24	1.4	2426	4	US-09-352-616A-470	Sequence 470, App
755	24	1.4	1781	1	US-08-314-362-2	Sequence 2, Appli	828	24	1.4	2426	4	US-09-534-638-3	Sequence 3, Appli
756	24	1.4	1781	1	US-08-433-010-2	Sequence 2, Appli	829	24	1.4	2480	4	US-09-605-785-332	Sequence 332, App
757	24	1.4	1781	1	US-08-245-295-4	Sequence 4, Appli	830	24	1.4	2507	4		



831	24	1.4	2507	4	US-09-439-313-332	Sequence 332, App	Sequence 332, App	c 904	24	1.4	3377	6	5198345-16	Patent No. 5198345
832	24	1.4	2507	4	US-09-352-616A-332	Sequence 332, App	Sequence 332, App	905	24	1.4	3437	3	US-08-704-711A-9	Sequence 9, Appl
833	24	1.4	2507	4	US-09-232-149A-332	Sequence 332, App	Sequence 332, App	906	24	1.4	3437	4	US-09-521-220-9	Sequence 9, Appl
834	24	1.4	2511	2	US-08-422-699A-8	Sequence 8, Appl	Sequence 8, Appl	907	24	1.4	3454	4	US-09-082-059-1	Sequence 1, Appl
835	24	1.4	2511	2	US-08-422-706B-8	Sequence 8, Appl	Sequence 8, Appl	908	24	1.4	3456	3	US-08-704-711A-8	Sequence 8, Appl
c 836	24	1.4	2559	2	US-08-724-774B-3	Sequence 3, Appl	Sequence 3, Appl	909	24	1.4	3456	4	US-09-521-220-8	Sequence 8, Appl
c 837	24	1.4	2559	3	US-09-089-593-3	Sequence 3, Appl	Sequence 3, Appl	910	24	1.4	3471	5	PCT-US93-00227-2	Sequence 2, Appl
c 838	24	1.4	2559	4	US-09-382-855-3	Sequence 3, Appl	Sequence 3, Appl	911	24	1.4	3487	4	US-09-303-069-24	Sequence 24, Appl
c 839	24	1.4	2559	4	US-09-183-714B-3	Sequence 3, Appl	Sequence 3, Appl	912	24	1.4	3507	1	US-08-832-883-67	Sequence 67, Appl
c 840	24	1.4	2559	4	US-09-642-281-3	Sequence 3, Appl	Sequence 3, Appl	913	24	1.4	3507	1	US-08-832-883-67	Sequence 67, Appl
841	24	1.4	2581	2	US-09-013-634-1	Sequence 1, Appl	Sequence 1, Appl	914	24	1.4	3507	2	US-08-832-877-67	Sequence 67, Appl
c 842	24	1.4	2605	2	US-08-680-395-4	Sequence 4, Appl	Sequence 4, Appl	915	24	1.4	3507	2	US-08-832-877-67	Sequence 67, Appl
843	24	1.4	2614	2	US-08-795-868-15	Sequence 15, Appl	Sequence 15, Appl	916	24	1.4	3603	4	US-09-499-884-11	Sequence 11, Appl
844	24	1.4	2614	4	US-09-303-069-15	Sequence 15, Appl	Sequence 15, Appl	917	24	1.4	3804	2	US-08-483-488-5	Sequence 5, Appl
845	24	1.4	2614	4	US-09-134-250-15	Sequence 15, Appl	Sequence 15, Appl	918	24	1.4	4066	4	US-09-367-750-1	Sequence 1, Appl
846	24	1.4	2628	1	US-08-143-219-1	Sequence 1, Appl	Sequence 1, Appl	919	24	1.4	4084	3	US-08-866-340-1	Sequence 1, Appl
847	24	1.4	2735	1	US-08-698-551-11	Sequence 11, Appl	Sequence 11, Appl	920	24	1.4	4137	4	US-09-499-964-2	Sequence 2, Appl
848	24	1.4	2735	2	US-08-602-228-11	Sequence 11, Appl	Sequence 11, Appl	921	24	1.4	4169	4	US-09-166-350-32	Sequence 32, Appl
849	24	1.4	2735	2	US-08-494-440B-11	Sequence 11, Appl	Sequence 11, Appl	922	24	1.4	4181	4	US-09-643-597-175	Sequence 175, App
850	24	1.4	2735	2	US-08-533-901B-11	Sequence 11, Appl	Sequence 11, Appl	923	24	1.4	4302	4	US-09-245-281-38	Sequence 38, Appl
851	24	1.4	2735	2	US-08-839-032A-11	Sequence 11, Appl	Sequence 11, Appl	924	24	1.4	4302	4	US-09-207-359B-38	Sequence 38, Appl
852	24	1.4	2735	2	US-08-839-031A-11	Sequence 11, Appl	Sequence 11, Appl	925	24	1.4	4421	2	US-08-257-963B-9	Sequence 9, Appl
853	24	1.4	2735	4	US-09-185-258C-11	Sequence 11, Appl	Sequence 11, Appl	926	24	1.4	4421	4	US-08-367-841A-9	Sequence 9, Appl
854	24	1.4	2735	5	PCT-US95-12724-11	Sequence 11, Appl	Sequence 11, Appl	927	24	1.4	4421	4	US-08-520-373D-6	Sequence 6, Appl
855	24	1.4	2738	2	US-08-795-868-17	Sequence 17, Appl	Sequence 17, Appl	928	24	1.4	4421	5	PCT-US95-07201-9	Sequence 9, Appl
856	24	1.4	2738	4	US-09-303-069-17	Sequence 17, Appl	Sequence 17, Appl	929	24	1.4	4517	4	US-09-140-804-9	Sequence 9, Appl
857	24	1.4	2738	4	US-09-134-250-17	Sequence 17, Appl	Sequence 17, Appl	930	24	1.4	4576	1	US-08-832-883-49	Sequence 49, Appl
858	24	1.4	2808	3	US-08-870-126-7	Sequence 7, Appl	Sequence 7, Appl	931	24	1.4	4576	2	US-08-832-877-49	Sequence 49, Appl
859	24	1.4	2808	4	US-09-445-247-7	Sequence 7, Appl	Sequence 7, Appl	932	24	1.4	4698	1	US-07-807-043B-5	Sequence 5, Appl
860	24	1.4	2817	1	US-08-655-836-4	Sequence 4, Appl	Sequence 4, Appl	933	24	1.4	4698	1	US-08-299-849B-5	Sequence 5, Appl
861	24	1.4	2817	2	US-09-020-753-4	Sequence 4, Appl	Sequence 4, Appl	934	24	1.4	4698	3	US-08-142-368A-5	Sequence 5, Appl
c 862	24	1.4	2839	4	US-09-061-702-1	Sequence 1, Appl	Sequence 1, Appl	935	24	1.4	4698	2	US-08-967-727-5	Sequence 5, Appl
863	24	1.4	2847	4	US-09-149-476-100	Sequence 100, App	Sequence 100, App	936	24	1.4	4698	4	US-08-037-230B-5	Sequence 5, Appl
864	24	1.4	2877	4	US-09-235-103-1	Sequence 1, Appl	Sequence 1, Appl	937	24	1.4	4749	1	US-08-452-259-1	Sequence 1, Appl
865	24	1.4	2904	4	US-09-403-618A-7	Sequence 7, Appl	Sequence 7, Appl	938	24	1.4	4749	5	PCT-US96-07336-1	Sequence 1, Appl
866	24	1.4	2909	6	5378464-1	Patent No. 5378464	Patent No. 5378464	939	24	1.4	4803	4	US-09-197-636-3	Sequence 3, Appl
867	24	1.4	3016	2	US-08-318-826A-7	Sequence 7, Appl	Sequence 7, Appl	940	24	1.4	4803	4	US-09-197-636-3	Sequence 3, Appl
868	24	1.4	3016	2	US-08-370-156-5	Sequence 5, Appl	Sequence 5, Appl	941	24	1.4	5232	3	US-09-212-971-3	Sequence 3, Appl
869	24	1.4	3016	3	US-08-814-095-5	Sequence 5, Appl	Sequence 5, Appl	942	24	1.4	5232	3	US-08-800-929A-3	Sequence 3, Appl
870	24	1.4	3024	6	5284931-1	Patent No. 5284931	Patent No. 5284931	943	24	1.4	5232	4	US-09-617-053A-3	Sequence 3, Appl
c 871	24	1.4	3046	1	US-08-726-725-1	Sequence 1, Appl	Sequence 1, Appl	944	24	1.4	5261	1	US-08-045-806-3	Sequence 3, Appl
872	24	1.4	3096	2	US-08-318-826A-6	Sequence 6, Appl	Sequence 6, Appl	945	24	1.4	5261	1	US-08-366-051B-3	Sequence 3, Appl
873	24	1.4	3096	2	US-08-370-156-3	Sequence 3, Appl	Sequence 3, Appl	946	24	1.4	5262	4	US-08-520-373D-5	Sequence 5, Appl
874	24	1.4	3096	3	US-08-814-095-3	Sequence 3, Appl	Sequence 3, Appl	947	24	1.4	5375	3	US-08-757-273-7	Sequence 7, Appl
875	24	1.4	3124	4	US-09-734-030-1	Sequence 1, Appl	Sequence 1, Appl	948	24	1.4	5555	1	US-08-484-438-3	Sequence 3, Appl
876	24	1.4	3158	2	US-08-464-517-36	Sequence 36, Appl	Sequence 36, Appl	949	24	1.4	5769	1	US-08-652-971-1	Sequence 1, Appl
877	24	1.4	3158	2	US-08-246-361A-36	Sequence 36, Appl	Sequence 36, Appl	950	24	1.4	5769	2	US-08-991-258A-1	Sequence 1, Appl
878	24	1.4	3158	3	US-08-463-772-36	Sequence 36, Appl	Sequence 36, Appl	951	24	1.4	5769	2	US-08-769-399-1	Sequence 1, Appl
879	24	1.4	3171	4	US-09-079-812E-1	Sequence 1, Appl	Sequence 1, Appl	952	24	1.4	5769	3	US-08-991-953A-1	Sequence 1, Appl
880	24	1.4	3176	2	US-08-910-733-17	Sequence 17, Appl	Sequence 17, Appl	953	24	1.4	5835	4	US-09-033-333-3	Sequence 3, Appl
881	24	1.4	3176	2	US-08-910-884-17	Sequence 17, Appl	Sequence 17, Appl	954	24	1.4	5835	4	US-09-033-333-3	Sequence 3, Appl
882	24	1.4	3212	4	US-08-697-954-1	Sequence 1, Appl	Sequence 1, Appl	955	24	1.4	5835	4	US-09-614-495-3	Sequence 3, Appl
883	24	1.4	3223	2	US-08-620-694A-9	Sequence 9, Appl	Sequence 9, Appl	956	24	1.4	5836	1	US-08-380-916-1	Sequence 1, Appl
884	24	1.4	3223	3	US-09-022-255-9	Sequence 9, Appl	Sequence 9, Appl	957	24	1.4	5836	3	US-08-721-690-1	Sequence 1, Appl
885	24	1.4	3223	3	US-09-022-696-9	Sequence 9, Appl	Sequence 9, Appl	958	24	1.4	5836	3	US-08-891-581-1	Sequence 1, Appl
886	24	1.4	3223	3	US-08-978-773-3	Sequence 3, Appl	Sequence 3, Appl	959	24	1.4	5836	4	US-09-033-333-2	Sequence 2, Appl
887	24	1.4	3223	3	US-09-022-253-9	Sequence 9, Appl	Sequence 9, Appl	960	24	1.4	5836	4	US-09-033-333-2	Sequence 2, Appl
888	24	1.4	3223	3	US-09-022-260-9	Sequence 9, Appl	Sequence 9, Appl	961	24	1.4	5836	4	US-09-614-495-2	Sequence 2, Appl
889	24	1.4	3223	4	US-09-022-259-9	Sequence 9, Appl	Sequence 9, Appl	962	24	1.4	5933	4	US-09-383-630-1	Sequence 1, Appl
890	24	1.4	3223	4	US-09-022-257-9	Sequence 9, Appl	Sequence 9, Appl	963	24	1.4	5933	4	US-09-383-630-2	Sequence 2, Appl
c 891	24	1.4	3267	2	US-08-257-963B-12	Sequence 12, Appl	Sequence 12, Appl	964	24	1.4	6002	1	US-08-698-551-15	Sequence 15, Appl
c 892	24	1.4	3267	4	US-08-367-841A-12	Sequence 12, Appl	Sequence 12, Appl	965	24	1.4	6002	2	US-08-602-228-15	Sequence 15, Appl
c 893	24	1.4	3267	5	PCT-US95-07201-12	Sequence 12, Appl	Sequence 12, Appl	966	24	1.4	6002	2	US-08-839-032A-15	Sequence 15, Appl
894	24	1.4	3287	4	US-08-811-481-15	Sequence 15, Appl	Sequence 15, Appl	967	24	1.4	6002	4	US-09-185-258C-15	Sequence 15, Appl
895	24	1.4	3320	4	US-09-394-200-1	Sequence 1, Appl	Sequence 1, Appl	968	24	1.4	6140	4	US-09-605-785-536	Sequence 536, App
896	24	1.4	3323	2	US-08-422-699A-10	Sequence 10, Appl	Sequence 10, Appl	969	24	1.4	6140	4	US-09-439-313-536	Sequence 536, App
897	24	1.4	3323	2	US-08-422-706B-10	Sequence 10, Appl	Sequence 10, Appl	970	24	1.4	6340	1	US-08-187-785-3	Sequence 3, Appl
898	24	1.4	3350	2	US-08-663-566A-1	Sequence 1, Appl	Sequence 1, Appl	971	24	1.4	6769	1	US-08-480-784-20	Sequence 20, Appl
899	24	1.4	3350	2	US-08-023-610-1	Sequence 1, Appl	Sequence 1, Appl	972	24	1.4	6769	1	US-08-483-553-20	Sequence 20, Appl
900	24	1.4	3350	2	US-08-288-065A-1	Sequence 1, Appl	Sequence 1, Appl	973	24	1.4	6769	1	US-08-487-002-20	Sequence 20, Appl
901	24	1.4	3350	2	US-08-362-240A-1	Sequence 1, Appl	Sequence 1, Appl	974	24	1.4	6769	1	US-08-483-554B-20	Sequence 20, Appl
902	24	1.4	3350	5	PCT-US95-10245-1	Sequence 1, Appl	Sequence 1, Appl	975	24	1.4	6769	1	US-08-488-011B-20	Sequence 20, Appl
c 903	24	1.4	3373	1	US-08-273-411-2	Sequence 2, Appl	Sequence 2, Appl	976	24	1.4	6769	4	US-08-850-727-20	Sequence 20, Appl

977 24 1.4 6769 5 PCT-US95-10202-20 Sequence 20, Appl  
978 24 1.4 6769 5 PCT-US95-10203-20 Sequence 20, Appl  
979 24 1.4 6769 5 PCT-US95-10220-20 Sequence 20, Appl  
980 24 1.4 6792 4 US-09-374-454-20 Sequence 20, Appl  
c 981 24 1.4 7505 4 US-09-078-294-13 Sequence 13, Appl  
982 24 1.4 8083 4 US-09-383-630-4 Sequence 4, Appl  
983 24 1.4 8083 4 US-09-383-630-5 Sequence 5, Appl  
c 984 24 1.4 8174 1 US-07-914-281-5 Sequence 5, Appl  
c 985 24 1.4 8174 1 US-08-393-246-5 Sequence 5, Appl  
c 986 24 1.4 8174 1 US-08-525-058A-5 Sequence 5, Appl  
c 987 24 1.4 8174 2 US-08-896-731-5 Sequence 5, Appl  
c 988 24 1.4 8174 4 US-09-042-531-5 Sequence 5, Appl  
c 989 24 1.4 8174 5 PCT-US91-00899-3 Sequence 3, Appl  
990 24 1.4 8353 3 US-08-611-587-1 Sequence 1, Appl  
991 24 1.4 8453 4 US-09-167-681-45 Sequence 45, Appl  
992 24 1.4 8779 2 US-08-750-703-4 Sequence 4, Appl  
c 993 24 1.4 9704 4 US-09-814-951A-3 Sequence 3, Appl  
994 24 1.4 9775 4 US-08-977-171-1 Sequence 1, Appl  
995 24 1.4 11531 1 US-08-068-945A-1 Sequence 1, Appl  
996 24 1.4 11531 1 US-08-442-806-1 Sequence 1, Appl  
c 997 24 1.4 11811 4 US-09-078-294-7 Sequence 7, Appl  
c 998 24 1.4 12047 2 US-09-022-461-1 Sequence 1, Appl  
c 999 24 1.4 12047 4 US-09-033-556-3 Sequence 3, Appl  
1000 24 1.4 13104 4 US-08-256-799-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1  
5508199-1  
; Patent No. 5508199  
; APPLICANT: GONZALES, FRANK J.; HARDWICK, JAMES P.; GELBOIN, HARRY V.; MEYER, URS A.  
; TITLE OF INVENTION: P450DB1 CLONES FOR IDENTIFYING HUMANS  
; WITH GENETIC DEFECT IN DRUG METABOLISM  
; NUMBER OF SEQUENCES: 11  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/242,158  
; FILING DATE: 13-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 845,507  
; FILING DATE: 27-FEB-1992  
; APPLICATION NUMBER: 292,815  
; FILING DATE: 03-JAN-1989  
; SEQ ID NO: 1:  
; LENGTH: 180

5508199-1  
Query Match 3.6%; Score 61; DB 6; Length 180;  
Best Local Similarity 100.0%; Pred. No. 1.2e-16;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTATAGTGCCCATCTCCTGCTCCTGGTG 1679  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTATAGTGCCCATCTCCTGCTCCTGGTG 60  
QY 1680 G 1680  
Db 61 G 61

RESULT 2  
US-08-145-658D-13  
; Sequence 13, Application US/08145658D  
; Patent No. 5981174  
; GENERAL INFORMATION:  
; APPLICANT: Wolf, Charles R.  
; APPLICANT: Miles, John S.  
; APPLICANT: Spurr, Nigel K.  
; APPLICANT: Gough, Alan C.  
; TITLE OF INVENTION: GENETIC ASSAY  
; NUMBER OF SEQUENCES: 25  
QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTATAGTGCCCATCTCCTGCTCCTGGTG 1679  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTATAGTGCCCATCTCCTGCTCCTGGTG 60  
QY 1680 G 1680  
Db 61 G 61

RESULT 3  
US-08-145-658D-22  
; Sequence 22, Application US/08145658D  
; Patent No. 5981174  
; GENERAL INFORMATION:  
; APPLICANT: Wolf, Charles R.  
; APPLICANT: Miles, John S.  
; APPLICANT: Spurr, Nigel K.  
; APPLICANT: Gough, Alan C.  
; TITLE OF INVENTION: GENETIC ASSAY  
; NUMBER OF SEQUENCES: 25  
QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTATAGTGCCCATCTCCTGCTCCTGGTG 1679  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTATAGTGCCCATCTCCTGCTCCTGGTG 60  
QY 1680 G 1680  
Db 61 G 61

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP  
; STREET: 2101 L Street N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/145,658D  
; FILING DATE: 04-NOV-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9001181.8  
; FILING DATE: 18-JAN-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/732,223  
; FILING DATE: 18-JUL-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brady, Jr., James W.  
; REGISTRATION NUMBER: 32,115  
; REFERENCE/DOCKET NUMBER: E8280.017  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-785-9700  
; TELEFAX: 202-887-0689  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1566 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; US-08-145-658D-13  
Query Match 3.6%; Score 61; DB 2; Length 1566;  
Best Local Similarity 100.0%; Pred. No. 1e-16;  
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTATAGTGCCCATCTCCTGCTCCTGGTG 1679  
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTATAGTGCCCATCTCCTGCTCCTGGTG 60  
QY 1680 G 1680  
Db 61 G 61  
RESULT 3  
US-08-145-658D-22  
; Sequence 22, Application US/08145658D  
; Patent No. 5981174  
; GENERAL INFORMATION:  
; APPLICANT: Wolf, Charles R.  
; APPLICANT: Miles, John S.  
; APPLICANT: Spurr, Nigel K.  
; APPLICANT: Gough, Alan C.  
; TITLE OF INVENTION: GENETIC ASSAY  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20037

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,658D
; FILING DATE: 04-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9001181.8
; FILING DATE: 18-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/732,223
; FILING DATE: 18-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, Jr., James W.
; REGISTRATION NUMBER: 32,115
; REFERENCE/DOCKET NUMBER: E8280.017
; TELEPHONE: 202-785-9700
; TELEFAX: 202-887-0689
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
; US-08-145-658D-22
;
; Query Match 3.6%; Score 61; DB 2; Length 1566;
; Best Local Similarity 100.0%; Pred. No. le-16;
; Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1620 ATGGGGCTAGAACGACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 1679
; |
; Db 1 ATGGGGCTAGAACGACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 60
;
; QY 1680 G 1680
; |
; Db 61 G 61
;
; RESULT 4
; US-08-145-658D-20
; Sequence 20, Application US/08145658D
; Patent No. 5981174
; GENERAL INFORMATION:
; APPLICANT: Wolf, Charles R.
; APPLICANT: Miles, John S.
; APPLICANT: Spurr, Nigel K.
; APPLICANT: Gough, Alan C.
; TITLE OF INVENTION: GENETIC ASSAY
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP
; STREET: 2101 L Street N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,658D
; FILING DATE: 04-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9001181.8
; FILING DATE: 18-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/732,223
; FILING DATE: 18-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, Jr., James W.
; REGISTRATION NUMBER: 32,115
; REFERENCE/DOCKET NUMBER: E8280.017
; TELEPHONE: 202-785-9700
; TELEFAX: 202-887-0689
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
; US-08-145-658D-22
;
; Query Match 3.6%; Score 61; DB 2; Length 1566;
; Best Local Similarity 100.0%; Pred. No. le-16;
; Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1620 ATGGGGCTAGAACGACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 1679
; |
; Db 1 ATGGGGCTAGAACGACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 60
;
; QY 1680 G 1680
; |
; Db 61 G 61
;
; RESULT 5
; US-08-145-658D-21
; Sequence 21, Application US/08145658D
; Patent No. 5981174
; GENERAL INFORMATION:
; APPLICANT: Wolf, Charles R.
; APPLICANT: Miles, John S.
; APPLICANT: Spurr, Nigel K.
; APPLICANT: Gough, Alan C.
; TITLE OF INVENTION: GENETIC ASSAY
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP
; STREET: 2101 L Street N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,658D
; FILING DATE: 04-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9001181.8
; FILING DATE: 18-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/732,223
; FILING DATE: 18-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, Jr., James W.
; REGISTRATION NUMBER: 32,115
; REFERENCE/DOCKET NUMBER: E8280.017
; TELEPHONE: 202-785-9700
; TELEFAX: 202-887-0689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
; US-08-145-658D-20
;
; Query Match 3.6%; Score 61; DB 2; Length 1568;
; Best Local Similarity 100.0%; Pred. No. le-16;
; Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1620 ATGGGGCTAGAACGACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 1679
; |
; Db 1 ATGGGGCTAGAACGACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 60
;
; QY 1680 G 1680
; |
; Db 61 G 61
;
; RESULT 5
; US-08-145-658D-21
; Sequence 21, Application US/08145658D
; Patent No. 5981174
; GENERAL INFORMATION:
; APPLICANT: Wolf, Charles R.
; APPLICANT: Miles, John S.
; APPLICANT: Spurr, Nigel K.
; APPLICANT: Gough, Alan C.
; TITLE OF INVENTION: GENETIC ASSAY
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP
; STREET: 2101 L Street N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,658D
; FILING DATE: 04-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9001181.8
; FILING DATE: 18-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/732,223
; FILING DATE: 18-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, Jr., James W.
; REGISTRATION NUMBER: 32,115
; REFERENCE/DOCKET NUMBER: E8280.017
; TELEPHONE: 202-785-9700
; TELEFAX: 202-887-0689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
; US-08-145-658D-20
```

```
;
; FILING DATE: 04-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9001181.8
; FILING DATE: 18-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/732,223
; FILING DATE: 18-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, Jr., James W.
; REGISTRATION NUMBER: 32,115
; REFERENCE/DOCKET NUMBER: E8280.017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-785-9700
; TELEFAX: 202-887-0689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
; US-08-145-658D-20
;
; Query Match 3.6%; Score 61; DB 2; Length 1568;
; Best Local Similarity 100.0%; Pred. No. le-16;
; Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1620 ATGGGGCTAGAACGACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 1679
; |
; Db 1 ATGGGGCTAGAACGACACTGGTGGCCCTGGCGTGATAGTGCCCATCTTCCTGCTCCTGGTG 60
;
; QY 1680 G 1680
; |
; Db 61 G 61
;
; RESULT 5
; US-08-145-658D-21
; Sequence 21, Application US/08145658D
; Patent No. 5981174
; GENERAL INFORMATION:
; APPLICANT: Wolf, Charles R.
; APPLICANT: Miles, John S.
; APPLICANT: Spurr, Nigel K.
; APPLICANT: Gough, Alan C.
; TITLE OF INVENTION: GENETIC ASSAY
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DICKSTEIN, SHAPIRO, MORIN & OSHINSKY LLP
; STREET: 2101 L Street N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/145,658D
; FILING DATE: 04-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9001181.8
; FILING DATE: 18-JAN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/732,223
; FILING DATE: 18-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, Jr., James W.
; REGISTRATION NUMBER: 32,115
; REFERENCE/DOCKET NUMBER: E8280.017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-785-9700
; TELEFAX: 202-887-0689
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
; US-08-145-658D-20
```

```
; FILING DATE: 18-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brady, Jr., James W.
; REGISTRATION NUMBER: 32,115
; REFERENCE/DOCKET NUMBER: E8280.017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-785-9700
; TELEFAX: 202-887-0689
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-145-658D-21

Query Match          3.6%; Score 61; DB 2; Length 1571;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1620 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCCTGCTCCTGGTG 1679
|
Db 1 ATGGGGCTAGAACACTGGTGGCCCTGGCGGTGATAGTGGCCATCTTCCTGCTCCTGGTG 60

QY 1680 G 1680
|
Db 61 G 61

RESULT 6
US-09-797-906-3
; Sequence 3, Application US/09797906
; Patent No. 6329188
; GENERAL INFORMATION:
; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL001151CIP
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 84495
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84495)
; OTHER INFORMATION: n = A,T,C or G
US-09-797-906-3

Query Match          2.5%; Score 42; DB 4; Length 84495;
Best Local Similarity 100.0%; Pred. No. 7e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 CACCTGTATATCCAGCTACTTAGGAGCTGAGGCGAGGAGAAAT 259
|
Db 45792 CACCTGTATATCCAGCTACTTAGGAGCTGAGGCGAGGAGAAAT 45833

RESULT 7
US-08-395-800A-7
; Sequence 7, Application US/08395800A
; Patent No. 5807732
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B
```

```
; APPLICANT: LENNON, GREGORY
; APPLICANT: ROQUIER, SYLVIE
; APPLICANT: GIORGI, DOMINIQUE
; APPLICANT: KELLY, ROBERT J
; TITLE OF INVENTION: GDP-L-FUCOSE: BETA-D-GALACTOSIDE
; TITLE OF INVENTION: 2-ALPHA-L-FUCOSYLTRANSFERASES, DNA SEQUENCES ENCODING THE
; TITLE OF INVENTION: SAME, METHOD FOR PRODUCING THE SAME AND A METHOD OF
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,800A
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2115 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 64..1092
US-08-395-800A-7

Query Match          2.4%; Score 41; DB 1; Length 2115;
Best Local Similarity 100.0%; Pred. No. 2.4e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAGACTAGCTGGCCCAACATGGTGAAACCC 176
|
Db 1316 GTCAGGAGTTCAGACTAGCTGGCCCAACATGGTGAAACCC 1356

RESULT 8
US-09-085-199B-44
; Sequence 44, Application US/09085199B
; Patent No. 6235879
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael R.
; APPLICANT: Hackam, Abigail
; APPLICANT: Huq, A.H.M. Mahbubul
; APPLICANT: Chopra, Vikramjit Singh
; APPLICANT: Kalchman, Michael
; TITLE OF INVENTION: Apoptosis Modulators That Interact with the
; TITLE OF INVENTION: Huntington's Disease Gene
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Oppedahl & Larson
; STREET: PO Box 5270
; CITY: Frisco
; STATE: CO
; COUNTRY: USA
; ZIP: 80443-5270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Kb storage
; COMPUTER: IBM Compatible
```

```

; OPERATING SYSTEM: MS DOS 5.0
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,199B
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Larson, Marina T.
; REGISTRATION NUMBER: 32038
; REFERENCE/DOCKET NUMBER: UBC.P-013US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (970) 668-2050
; TELEFAX: (970) 668-2052
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3715
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: genomic DNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; FEATURE:
; OTHER INFORMATION: exon 29 and partial cds of HIP1
US-09-085-199B-44

Query Match 2.4%; Score 41; DB 4; Length 3715;
Best Local Similarity 100.0%; Pred. No. 2.3e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCTGCCAACATGTTGAAACCC 176
|||||
Db 3588 GTCAGGAGTTCAAGACTAGCTGCCAACATGTTGAAACCC 3628

RESULT 9
US-09-754-250-3
; Sequence 3, Application US/09754250
; Patent No. 6376225
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001063
; CURRENT APPLICATION NUMBER: US/09/754,250
; CURRENT FILING DATE: 2001-01-05
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(111282)
; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match 2.4%; Score 41; DB 4; Length 111282;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 TCCGACCTACTTAGGAGGCTGAGGCAGGAGAAATTCGTTGAA 267
|||||
Db 3839 TCCGACCTACTTAGGAGGCTGAGGCAGGAGAAATTCGTTGAA 3879

RESULT 10
US-07-869-933-31
; Sequence 31, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11298 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; ORGANISM: homo sapien
; STRAIN: FCRI beta
US-07-869-933-31

Query Match 2.3%; Score 38; DB 1; Length 11298;
Best Local Similarity 100.0%; Pred. No. 3.9e-07;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 CAGCTACTTAGGAGGCTGAGGCAGGAGAAATTCGTTGAA 267
|||||
Db 6813 CAGCTACTTAGGAGGCTGAGGCAGGAGAAATTCGTTGAA 6850

RESULT 11
US-08-201-879A-2
; Sequence 2, Application US/08201879A
; Patent No. 5807988
; GENERAL INFORMATION:
; APPLICANT: KINET, Jean-Pierre
; APPLICANT: JOUVIN, Marie-Helene
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN E
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/201,879A  
FILING DATE: 24-FEB-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/869,933  
FILING DATE: 16-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/03419  
FILING DATE: 16-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BENT, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40399/234/NIHD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11298 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
STRAIN: FcRI beta  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(456..511, 1381..1510, 2026..2160, 4475..4531,  
LOCATION: 5079..5237, 5640..5738, 7224..7319)  
US-08-201-879A-2

Query Match 2.3%; Score 38; DB 1; Length 11298;

Best Local Similarity 100.0%; Pred. No. 3.9e-07;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 CAGCTACTTAGGAGGCTGAGGCAGGAGAAATTCCTTGA 267  
|||||  
Db 6813 CAGCTACTTAGGAGGCTGAGGCAGGAGAAATTCCTTGA 6850

## RESULT 12

US-09-103-663-31  
Sequence 31, Application US/09103663D  
Patent No. 6171803

GENERAL INFORMATION:  
APPLICANT: Kinet et al.  
TITLE OF INVENTION: Isolation, characterization, and use of the human beta  
TITLE OF INVENTION: subunit of the high affinity receptor for  
TITLE OF INVENTION: immunoglobulin E.  
FILE REFERENCE: 50490

CURRENT APPLICATION NUMBER: US/09/103,663D  
CURRENT FILING DATE: 1998-06-23

EARLIER APPLICATION NUMBER: 07/869,933  
EARLIER FILING DATE: 1992-04-16

NUMBER OF SEQ ID NOS: 35  
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 31

LENGTH: 11298

TYPE: DNA

ORGANISM: Homo sapiens

US-09-103-663-31

Query Match 2.3%; Score 38; DB 4; Length 11298;

Best Local Similarity 100.0%; Pred. No. 3.9e-07;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 230 CAGCTACTTAGGAGGCTGAGGCAGGAGAAATTCCTTGA 267  
|||||  
Db 6813 CAGCTACTTAGGAGGCTGAGGCAGGAGAAATTCCTTGA 6850

## RESULT 13

US-09-741-150-3/c

Sequence 3, Application US/09741150

Patent No. 6436689

GENERAL INFORMATION:

APPLICANT: GUEGLER, Karl et al

TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,

TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: CL000968

CURRENT APPLICATION NUMBER: US/09/741,150

CURRENT FILING DATE: 2000-12-21

NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 112132

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(112132)

OTHER INFORMATION: n = A,T,C or G

US-09-741-150-3

Query Match 2.3%; Score 38; DB 4; Length 112132;

Best Local Similarity 100.0%; Pred. No. 3.2e-07;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 GTGTGGCACACACCTGTAATCCAGCTACTTAGGAGG 244  
|||||

Db 80384 GTGTGGCACACACCTGTAATCCAGCTACTTAGGAGG 80347

## RESULT 14

US-08-257-963B-10

Sequence 10, Application US/08257963B

Patent No. 5840686

GENERAL INFORMATION:

APPLICANT: Chader, Gerald J.; Becerra, S.

APPLICANT: Patricia; Schwartz, Joan P.;

APPLICANT: Taniwaki, Takayuki

TITLE OF INVENTION: PIGMENT EPITHELIUM

TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL

TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING

TITLE OF INVENTION: AND EXPRESSING THE PROTEIN

NUMBER OF SEQUENCES: 42

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan & Finnegan

STREET: 345 Park Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/257,963B

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/952,796

FILING DATE: 24-SEPT-1992

ATTORNEY/AGENT INFORMATION:

NAME: DOROTHY R. AUTH

REGISTRATION NUMBER: 36434

REFERENCE/DOCKET NUMBER: 20264126US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7210 Base Pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Double  
TOPOLOGY: Unknown  
MOLECULE TYPE: Genomic DNA  
ORGANISM: Human  
IMMEDIATE SOURCE:  
LIBRARY: DASH II  
FEATURE:  
NAME/KEY: JT106  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: 7.2 kb No. 5840686 1 fragments  
US-08-257-963B-10

Query Match 2.2%; Score 37; DB 2; Length 7210;  
Best Local Similarity 100.0%; Pred. No. 1e-06;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
US-08-257-963B-10

QY 75 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 111  
|||||  
DB 3065 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 3101  
|||||

## RESULT 15

US-08-367-841A-10  
Sequence 10, Application US/08367841A  
Patent No. 6319687  
GENERAL INFORMATION:  
APPLICANT: Chader, Gerald J.; Rodriguez,  
APPLICANT: Ignacio R.; Mazuruk, Krzysztof;  
APPLICANT: Tombran-Tink, Joyce  
TITLE OF INVENTION: PIGMENT EPITHELIUM  
TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan  
STREET: 345 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/367,841A  
FILING DATE: 30-DEC-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,963  
FILING DATE: 07-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/952,796  
FILING DATE: 24-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36434  
REFERENCE/DOCKET NUMBER: 20264126US2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7210 Base Pairs  
TYPE: Nucleic Acid

STRANDEDNESS: Double  
TOPOLOGY: Unknown  
MOLECULE TYPE: Genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: Human  
IMMEDIATE SOURCE:  
LIBRARY: DASH II  
FEATURE:  
NAME/KEY: JT6A  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: 7.0 kb No. 6319687 1-No. 6319687  
fragment; Derived from human placental  
US-08-367-841A-10

Query Match 2.2%; Score 37; DB 4; Length 7210;  
Best Local Similarity 100.0%; Pred. No. 1e-06;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
US-08-367-841A-10

QY 75 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 111  
|||||  
DB 3065 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 3101  
|||||

## RESULT 16

PCT-US95-07201-10  
Sequence 10, Application PC/TUS9507201  
GENERAL INFORMATION:  
APPLICANT: Chader, Gerald J.; Becerra, Sofia  
APPLICANT: Patricia; Schwartz, Joan P.;  
APPLICANT: Taniwaki, Takayuki  
TITLE OF INVENTION: PIGMENT EPITHELIUM  
TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07201  
FILING DATE: 06-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/367,841  
FILING DATE: 30-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/257,963  
FILING DATE: 07-JUN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/952,796  
FILING DATE: 24-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36434  
REFERENCE/DOCKET NUMBER: 20264126PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7210 Base Pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Double

```

; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Human
; IMMEDIATE SOURCE:
; LIBRARY: DASH II
; FEATURE:
; NAME/KEY: JT6A
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 7.0 kb Not 1-Not
; OTHER INFORMATION: fragment; Derived from human placental
; OTHER INFORMATION: genomic DNA; also referred to as JT106
PCT-US95-07201-10

Query Match          2.2%; Score 37; DB 5; Length 7210;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 111
|||||
DB 3065 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 3101

RESULT 17
US-08-520-373D-4
; Sequence 4, Application US/08520373D
; Patent No. 6451763
; GENERAL INFORMATION:
; APPLICANT: Tombran-Tink, Joyce
; APPLICANT: Steele, Fintan R
; APPLICANT: Chader, Gerald J
; APPLICANT: Becerra, Sofia P
; APPLICANT: Johnson, Lincoln V
; APPLICANT: Rodriguez, Ignacio R
; TITLE OF INVENTION: RETINAL PIGMENTED EPITHELIUM DERIVED NEUROTROPIC FACTOR
; FILE REFERENCE: 2026-4203US1
; CURRENT APPLICATION NUMBER: US/08/520,373D
; CURRENT FILING DATE: 1995-08-29
; PRIOR FILING DATE: 08/377,710
; PRIOR FILING DATE: 1995-01-25
; PRIOR APPLICATION NUMBER: 08/279,979
; PRIOR FILING DATE: 1994-07-25
; PRIOR APPLICATION NUMBER: 07/894,215
; PRIOR FILING DATE: 1992-06-04
; PRIOR APPLICATION NUMBER: 07/952,796
; PRIOR FILING DATE: 1992-09-24
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 14581
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: mRNA: 6683; EXON: 6683-6790; EXON 11584-11675;
; OTHER INFORMATION: EXON: 14539-14581; INTRON: 6791-11583; INTRON:
; OTHER INFORMATION: 11676-14538; CDS: 11584-11675; 14539-14580
US-08-520-373D-4

Query Match          2.2%; Score 37; DB 4; Length 14581;
Best Local Similarity 100.0%; Pred. No. 9.9e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 111
|||||
DB 3064 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 3100

RESULT 18
US-08-367-841A-43
; Sequence 43, Application US/08367841A
; Patent No. 6319687
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Rodriguez,
; APPLICANT: Tombran-Tink, Joyce
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.

```

```

; APPLICANT: Chader, Gerald J.; Rodriguez,
; APPLICANT: Ignacio R.; Mazuruk, Krzysztof;
; APPLICANT: Tombran-Tink, Joyce
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/367,841A
; FILING DATE: 30-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22481 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: PI-147
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: full length genomic
; OTHER INFORMATION: sequence for PEDF plus flanking sequences.
US-08-367-841A-43

Query Match          2.2%; Score 37; DB 4; Length 22481;
Best Local Similarity 100.0%; Pred. No. 9.6e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 111
|||||
DB 3057 GCGGTGGCTCATGCTATATCCAGCAGCTTTGGGAG 3093

RESULT 19
PCT-US95-07201-43
; Sequence 43, Application PC/TUS9507201
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, Sofia
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.

```



STREET: 345 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/07201  
FILING DATE: 06-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/367,841  
FILING DATE: 30-DEC-1994  
PRIOR APPLICATION DATA: 08/257,963  
FILING DATE: 07-JUN-1994  
APPLICATION NUMBER: 07/952,796  
FILING DATE: 24-SEP-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: DOROTHY R. AUTH  
REGISTRATION NUMBER: 36434  
REFERENCE/DOCKET NUMBER: 20264126PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22481 Base Pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Double  
TOPOLOGY: Unknown  
MOLECULE TYPE: Genomic DNA  
FEATURE:  
NAME/KEY: Pl-147  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: full length genomic  
OTHER INFORMATION: sequence for PEDF plus flanking sequences.  
PCT-US95-07201-43

Query Match 2.2%; Score 37; DB 5; Length 22481;  
Best Local Similarity 100.0%; Pred. No. 9.6e-07;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GCGGTGGCTCATGCTATATCCAGCACCTTTGGGAG 111  
|||||  
Db 3057 GCGGTGGCTCATGCTATATCCAGCACCTTTGGGAG 3093

RESULT 20  
US-09-875-223-2  
Sequence 2, Application US/09875223  
Patent No. 6391850  
GENERAL INFORMATION:  
APPLICANT: No. 6391850thwestern University  
APPLICANT: No. 63918501 Bouck  
APPLICANT: David Dawson  
APPLICANT: Paul Gillis  
TITLE OF INVENTION: Methods and Compositions for Inhibiting Angiogenesis  
FILE REFERENCE: 0290-2303  
CURRENT APPLICATION NUMBER: US/09/875,223  
CURRENT FILING DATE: 2001-06-06  
PRIOR APPLICATION NUMBER: US 09/122,079  
PRIOR FILING DATE: 1998-07-23  
PRIOR APPLICATION NUMBER: PCT/US98/15228  
PRIOR FILING DATE: 1998-07-23  
PRIOR APPLICATION NUMBER: US 08/899,304  
PRIOR FILING DATE: 1997-07-23

NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 22484  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: Unsure  
LOCATION: 1...22484  
OTHER INFORMATION: "n" means either a, c, t, or g  
US-09-875-223-2

Query Match 2.2%; Score 37; DB 4; Length 22484;  
Best Local Similarity 100.0%; Pred. No. 9.6e-07;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GCGGTGGCTCATGCTATATCCAGCACCTTTGGGAG 111  
|||||  
Db 3057 GCGGTGGCTCATGCTATATCCAGCACCTTTGGGAG 3093

RESULT 21  
US-09-078-294-4/C  
Sequence 4, Application US/09078294  
Patent No. 6265211  
GENERAL INFORMATION:  
APPLICANT: Choo, Kong-Hong Andy  
APPLICANT: Du Sart, Desiree  
APPLICANT: Cancilla, Michael R.  
TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
FILE REFERENCE: Davies Col  
CURRENT APPLICATION NUMBER: US/09/078,294  
CURRENT FILING DATE: 1998-05-13  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 80246  
TYPE: DNA  
ORGANISM: Nucleotide sequence of NC-contig  
US-09-078-294-4

Query Match 2.2%; Score 37; DB 4; Length 80246;  
Best Local Similarity 100.0%; Pred. No. 8.7e-07;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 CTGAGTGGGTGGATCACCTGAAGTCAGGAGTTCAAG 149  
|||||  
Db 20206 CTGAGTGGGTGGATCACCTGAAGTCAGGAGTTCAAG 20170

RESULT 22  
US-09-078-294-3/C  
Sequence 3, Application US/09078294  
Patent No. 6265211  
GENERAL INFORMATION:  
APPLICANT: Choo, Kong-Hong Andy  
APPLICANT: Du Sart, Desiree  
APPLICANT: Cancilla, Michael R.  
TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
FILE REFERENCE: Davies Col  
CURRENT APPLICATION NUMBER: US/09/078,294  
CURRENT FILING DATE: 1998-05-13  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 80595  
TYPE: DNA  
ORGANISM: Nucleotide sequence of HC-contig  
US-09-078-294-3

Query Match 2.2%; Score 37; DB 4; Length 80595;  
Best Local Similarity 100.0%; Pred. No. 8.7e-07;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 113 CTGAGTGGGTGATCACTGAAGTCAGGAGTCAAG 149  
|||||  
Db 20469 CTGAGTGGGTGATCACTGAAGTCAGGAGTCAAG 20433

## RESULT 23

US-09-345-882-1/c  
; Sequence 1, Application US/09345882  
; Patent No. 6399373  
; GENERAL INFORMATION:  
; APPLICANT: Bougueleret, Lydie  
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)  
; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.  
; FILE REFERENCE: GENSET.031A  
; CURRENT APPLICATION NUMBER: US/09/345,882  
; CURRENT FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: US 60/091,315  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/111,909  
; PRIOR FILING DATE: 1998-12-10  
; NUMBER OF SEQ ID NOS: 140  
; SOFTWARE: Patent.pm  
; SEQ ID NO 1  
; LENGTH: 162450  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 72794  
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 88073  
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 90842  
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 93714  
; OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 97122  
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 97152  
; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 99098  
; OTHER INFORMATION: 5-130-257 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 99117  
; OTHER INFORMATION: 5-130-276 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 103806  
; OTHER INFORMATION: 5-131-395 : polymorphic base A or T  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 106940  
; OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 108106  
; OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A  
; FEATURE:  
; NAME/KEY: allele

; LOCATION: 108149  
; OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 108308  
; OTHER INFORMATION: 5-135-357 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 108471  
; OTHER INFORMATION: 5-136-174 : polymorphic base C or T  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 134134  
; OTHER INFORMATION: 5-140-120 : polymorphic base C or T  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 134362  
; OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 134374  
; OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 146328  
; OTHER INFORMATION: 5-143-84 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 146345  
; OTHER INFORMATION: 5-143-101 : polymorphic base A or C  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 150329  
; OTHER INFORMATION: 5-145-24 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 160031  
; OTHER INFORMATION: 5-148-352 : polymorphic base G or T  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 72771..72817  
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 72771..72817  
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 88050..88096  
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 88050..88096  
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 90819..90865  
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 90819..90865  
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 93690..93736  
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 93690..93736  
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 97099..97145

OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 97099..97145  
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 97130..97177  
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 97130..97177  
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 99075..99121  
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 99075..99121  
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 99094..99140  
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 103783..103828  
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 103783..103828  
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 106918..106966  
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 106918..106966  
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 108084..108130  
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 108084..108130  
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 108127..108177  
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39  
FEATURE: allele  
NAME/KEY: allele  
LOCATION: 108127..108177  
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60  
FEATURE:

Query Match 2.2%; Score 37; DB 4; Length 162450;  
Best Local Similarity 100.0%; Pred. No. 8.3e-07;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 GCGGTGGCTCATCCCTATATCCAGCAGCTTTGGGAG 111  
|||||

Db 51622 GCGGTGGCTCATCCCTATATCCAGCAGCTTTGGGAG 51586

RESULT 24

US-09-128-155-16  
; Sequence 16, Application US/09128155  
; Patent No. 6117654  
; GENERAL INFORMATION:  
; APPLICANT: Bao, Yang  
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
; FILE REFERENCE: 09404/052001  
; CURRENT APPLICATION NUMBER: US/09/128,155  
; CURRENT FILING DATE: 1998-08-03  
; EARLIER APPLICATION NUMBER: US 60/091,650  
; EARLIER FILING DATE: 1998-07-02  
; EARLIER APPLICATION NUMBER: US 60/054,646  
; EARLIER FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 152331  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(152331)  
; OTHER INFORMATION: n - A,T,C or G  
US-09-128-155-16

Query Match 2.1%; Score 36; DB 3; Length 152331;  
Best Local Similarity 100.0%; Pred. No. 2.2e-06;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 AGGTGGTGGATCACCTGAAGTCAGGAGTTCAAGAC 151  
|||||

Db 151040 AGGTGGTGGATCACCTGAAGTCAGGAGTTCAAGAC 151075

RESULT 25

US-09-426-290-1/c  
; Sequence 1, Application US/09426290  
; Patent No. 6410712  
; GENERAL INFORMATION:  
; APPLICANT: Berglind Ran Olafsdottir  
; APPLICANT: Jeffrey Gulcher  
; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE  
; FILE REFERENCE: 2345.2001-000  
; CURRENT APPLICATION NUMBER: US/09/426,290  
; CURRENT FILING DATE: 1999-10-25  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 168575  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (21181)..(21403)  
; NAME/KEY: CDS  
; LOCATION: (95252)..(95430)  
; NAME/KEY: CDS  
; LOCATION: (101753)..(101996)  
; NAME/KEY: CDS  
; LOCATION: (110324)..(110439)  
; NAME/KEY: CDS  
; LOCATION: (124058)..(124278)  
; NAME/KEY: CDS  
; LOCATION: (127009)..(127130)  
; NAME/KEY: CDS  
; LOCATION: (128910)..(129139)  
US-09-426-290-1

Query Match 2.1%; Score 36; DB 4; Length 168575;  
Best Local Similarity 100.0%; Pred. No. 2.2e-06;  
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 CGGTGGCTCATGCTATATAATCCAGCACTTTGGGAG 111  
|||||  
Db 65704 CGGTGGCTCATGCTATATAATCCAGCACTTTGGGAG 65669

RESULT 26  
US-09-305-384-5  
; Sequence 5, Application US/09305384  
; Patent No. 6242218  
; GENERAL INFORMATION:  
; APPLICANT: Treco, Douglas A.  
; APPLICANT: Heartlein, Michael W.  
; APPLICANT: Selden, Richard F.  
; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY  
; FILE REFERENCE: 07236/017001  
; CURRENT APPLICATION NUMBER: US/09/305,384  
; CURRENT FILING DATE: 1999-05-05  
; EARLIER APPLICATION NUMBER: US 60/084,649  
; EARLIER FILING DATE: 1998-05-07  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 6235  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-305-384-5

Query Match 2.1%; Score 35; DB 4; Length 6235;  
Best Local Similarity 100.0%; Pred. No. 7.3e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 AGACTAGCCTGGCCAAACATGGTGAACCCCTATCTC 182  
|||||  
Db 159 AGACTAGCCTGGCCAAACATGGTGAACCCCTATCTC 193

RESULT 27  
US-09-305-384-1  
; Sequence 1, Application US/09305384  
; Patent No. 6242218  
; GENERAL INFORMATION:  
; APPLICANT: Treco, Douglas A.  
; APPLICANT: Heartlein, Michael W.  
; APPLICANT: Selden, Richard F.  
; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY  
; FILE REFERENCE: 07236/017001  
; CURRENT APPLICATION NUMBER: US/09/305,384  
; CURRENT FILING DATE: 1999-05-05  
; EARLIER APPLICATION NUMBER: US 60/084,649  
; EARLIER FILING DATE: 1998-05-07  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 6679  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-305-384-1

Query Match 2.1%; Score 35; DB 4; Length 6679;  
Best Local Similarity 100.0%; Pred. No. 7.3e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 AGACTAGCCTGGCCAAACATGGTGAACCCCTATCTC 182  
|||||  
Db 178 AGACTAGCCTGGCCAAACATGGTGAACCCCTATCTC 212

RESULT 28  
US-09-738-884-3  
; Sequence 3, Application US/09738884  
; Patent No. 6391606  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al

; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000849  
; CURRENT APPLICATION NUMBER: US/09/738,884  
; CURRENT FILING DATE: 2000-12-18  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 13953  
; TYPE: DNA  
; ORGANISM: Human  
US-09-738-884-3

Query Match 2.1%; Score 35; DB 4; Length 13953;  
Best Local Similarity 100.0%; Pred. No. 6.9e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTATATCCAGCTACTTAGGAGGCTGAGGCAGG 254  
|||||  
Db 7286 CCTGTATATCCAGCTACTTAGGAGGCTGAGGCAGG 7320

RESULT 29  
US-09-729-995-3  
; Sequence 3, Application US/09729995  
; Patent No. 6426206  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL000904  
; CURRENT APPLICATION NUMBER: US/09/729,995  
; CURRENT FILING DATE: 2000-12-06  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 29629  
; TYPE: DNA  
; ORGANISM: Human  
US-09-729-995-3

Query Match 2.1%; Score 35; DB 4; Length 29629;  
Best Local Similarity 100.0%; Pred. No. 6.5e-06;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GGTGGCTCATGCTATATAATCCAGCACTTTGGGAG 111  
|||||  
Db 24605 GGTGGCTCATGCTATATAATCCAGCACTTTGGGAG 24639

RESULT 30  
US-09-735-934A-3  
; Sequence 3, Application US/09735934A  
; Patent No. 6372468  
; GENERAL INFORMATION:  
; APPLICANT: LI, Jiaxin et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL000851  
; CURRENT APPLICATION NUMBER: US/09/735,934A  
; CURRENT FILING DATE: 2000-12-14  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 43950  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-735-934A-3

Query Match 2.1%; Score 35; DB 4; Length 43950;

```
Best Local Similarity 100.0%; Pred. No. 6.3e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GTGGGCTCATGCTATAATCCAGCAGCACTTTGGGAG 111
      |||||
Db 8966 GTGGGCTCATGCTATAATCCAGCAGCACTTTGGGAG 9000

RESULT 31
US-09-750-580-1/c
; Sequence 1, Application US/09750580
; Patent No. 6455280
; GENERAL INFORMATION:
; APPLICANT: Yen, Frances
; APPLICANT: Denison, Blake
; APPLICANT: Bour, Barbara
; APPLICANT: Bihaïn, Bernard
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouguieret, Lydie
; APPLICANT: Ebbets-Reed, Dana
; APPLICANT: Salter-Cid, Luisa
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING NEOPLASTIC CELL GROWTH
; FILE REFERENCE: 89 US2.CIP
; CURRENT APPLICATION NUMBER: US/09/750,580
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 09/599,362
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: PCT/IB00/0101
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: PCT/IB99/02058
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: US 49/469/099
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 60/141,032
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 81001
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 10946..12946
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 12947..12958
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 13470..13526
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 13641..13752
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 14271..15968
; OTHER INFORMATION: exon 4
; NAME/KEY: misc.feature
; LOCATION: 15969..17969
; OTHER INFORMATION: 3'regulatory region
; NAME/KEY: allele
; LOCATION: 1239
; OTHER INFORMATION: 20-828-311 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 12347
; OTHER INFORMATION: 17-42-319 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 15241
; OTHER INFORMATION: 17-41-250 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 45461

; LOCATION: 42218
; OTHER INFORMATION: 20-841-149 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 45442
; OTHER INFORMATION: 20-842-115 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 77058
; OTHER INFORMATION: 20-853-415 : polymorphic base C or T
; NAME/KEY: primer_bind
; LOCATION: 929..949
; OTHER INFORMATION: 20-828.pu
; NAME/KEY: primer_bind
; LOCATION: 1357..1377
; OTHER INFORMATION: 20-828.rp complement
; NAME/KEY: primer_bind
; LOCATION: 12029..12050
; OTHER INFORMATION: 17-42.pu
; NAME/KEY: primer_bind
; LOCATION: 12581..12603
; OTHER INFORMATION: 17-42.rp complement
; NAME/KEY: primer_bind
; LOCATION: 14992..15012
; OTHER INFORMATION: 17-41.pu
; NAME/KEY: primer_bind
; LOCATION: 15460..15482
; OTHER INFORMATION: 17-41.rp complement
; NAME/KEY: primer_bind
; LOCATION: 42070..42090
; OTHER INFORMATION: 20-841.pu
; NAME/KEY: primer_bind
; LOCATION: 42572..42591
; OTHER INFORMATION: 20-841.rp complement
; NAME/KEY: primer_bind
; LOCATION: 45328..45347
; OTHER INFORMATION: 20-842.pu
; NAME/KEY: primer_bind
; LOCATION: 45863..45883
; OTHER INFORMATION: 20-842.rp complement
; NAME/KEY: primer_bind
; LOCATION: 76644..76664
; OTHER INFORMATION: 20-853.pu
; NAME/KEY: primer_bind
; LOCATION: 77166..77185
; OTHER INFORMATION: 20-853.rp complement
; NAME/KEY: primer_bind
; LOCATION: 1220..1238
; OTHER INFORMATION: 20-828-311.mis
; NAME/KEY: primer_bind
; LOCATION: 1240..1258
; OTHER INFORMATION: 20-828-311.mis complement
; NAME/KEY: primer_bind
; LOCATION: 12328..12346
; OTHER INFORMATION: 17-42-319.mis
; NAME/KEY: primer_bind
; LOCATION: 12348..12366
; OTHER INFORMATION: 17-42-319.mis complement
; NAME/KEY: primer_bind
; LOCATION: 15222..15240
; OTHER INFORMATION: 17-41-250.mis
; NAME/KEY: primer_bind
; LOCATION: 15242..15260
; OTHER INFORMATION: 17-41-250.mis complement
; NAME/KEY: primer_bind
; LOCATION: 42199..42217
; OTHER INFORMATION: 20-841-149.mis
; NAME/KEY: primer_bind
; LOCATION: 42219..42237
; OTHER INFORMATION: 20-841-149.mis complement
; NAME/KEY: primer_bind
; LOCATION: 45423..45441
; OTHER INFORMATION: 20-842-115.mis
; NAME/KEY: primer_bind
; LOCATION: 45443..45461
```

```
; OTHER INFORMATION: 20-842-115.mis complement
; NAME/KEY: primer_bind
; LOCATION: 77039..77057
; OTHER INFORMATION: 20-853-415.mis
; NAME/KEY: primer_bind
; LOCATION: 77059..77077
; OTHER INFORMATION: 20-853-415.mis complement
; NAME/KEY: misc_binding
; LOCATION: 1227..1251
; OTHER INFORMATION: 20-828-311.probe
; NAME/KEY: misc_binding
; LOCATION: 12335..12359
; OTHER INFORMATION: 17-42-319.probe
; NAME/KEY: misc_binding
; LOCATION: 15229..15253
; OTHER INFORMATION: 17-41-250.probe
; NAME/KEY: misc_binding
; LOCATION: 42206..42230
; OTHER INFORMATION: 20-841-149.probe
; NAME/KEY: misc_binding
; LOCATION: 45430..45454
; OTHER INFORMATION: 20-842-115.probe
; NAME/KEY: misc_binding
; LOCATION: 77046..77070
; OTHER INFORMATION: 20-853-415.probe
; US-09-750-580-1

Query Match          2.1%; Score 35; DB 4; Length 81001;
Best Local Similarity 100.0%; Pred. No. 6e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GTGGCTCATGCTATATATCCAGCAGCACTTTGGGAG 111
|||||
Db 3067 GGTGGCTCATGCTATATATCCAGCAGCACTTTGGGAG 3033

RESULT 32
US-09-798-096-10/c
; Sequence 10, Application US/09798096
; Patent No. 6399378
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; TITLE OF INVENTION: ANTISENSE MODULATION OF RECQL2 EXPRESSION
; FILE REFERENCE: RTS-0207
; CURRENT APPLICATION NUMBER: US/09/798,096
; CURRENT FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 10
; LENGTH: 99500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; US-09-798-096-10

Query Match          2.1%; Score 35; DB 4; Length 99500;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GTGGCTCATGCTATATATCCAGCAGCACTTTGGGAG 111
|||||
Db 69587 GGTGGCTCATGCTATATATCCAGCAGCACTTTGGGAG 69553

RESULT 33
US-09-305-384-6
; Sequence 6, Application US/09305384
; Patent No. 6242218
; GENERAL INFORMATION:
; APPLICANT: Treco, Douglas A.
; APPLICANT: Heartlein, Michael W.
; APPLICANT: Selden, Richard F.
; TITLE OF INVENTION: GENOMIC SEQUENCES FOR PROTEIN PRODUCTION AND DELIVERY
```

```
; FILE REFERENCE: 07236/017001
; CURRENT APPLICATION NUMBER: US/09/305,384
; CURRENT FILING DATE: 1999-05-05
; EARLIER APPLICATION NUMBER: US 60/084,649
; EARLIER FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 2834
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-305-384-6

Query Match          2.0%; Score 34; DB 4; Length 2834;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GTGGCTCATGCTATATATCCAGCAGCACTTTGGGAG 111
|||||
Db 2692 GTGGCTCATGCTATATATCCAGCAGCACTTTGGGAG 2725

RESULT 34
US-09-632-098-1
; Sequence 1, Application US/09632098
; Patent No. 6420154
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Baindur, Nand
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: MAMMALIAN ADHESION PROTEASE PEPTIDES
; FILE REFERENCE: 99-39
; CURRENT APPLICATION NUMBER: US/09/632,098
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 3431
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (37)...(2442)
; US-09-632-098-1

Query Match          2.0%; Score 34; DB 4; Length 3431;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GTGGCTCATGCTATATATCCAGCAGCACTTTGGGA 110
|||||
Db 3134 GTGGCTCATGCTATATATCCAGCAGCACTTTGGGA 3167

RESULT 35
US-09-632-098-3
; Sequence 3, Application US/09632098
; Patent No. 6420154
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Baindur, Nand
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: MAMMALIAN ADHESION PROTEASE PEPTIDES
; FILE REFERENCE: 99-39
; CURRENT APPLICATION NUMBER: US/09/632,098
; CURRENT FILING DATE: 2000-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 3468
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```

; NAME/KEY: CDS
; LOCATION: (37)...(2472)
US-09-632-098-3

Query Match      2.0%; Score 34; DB 4; Length 3468;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 77 GTGGCTCATGCGCTATATATCCAGCACTTTGGGA 110
|||||
Db 3171 GTGGCTCATGCGCTATATATCCAGCACTTTGGGA 3204

RESULT 36
US-08-884-324-9
; Sequence 9, Application US/08884324
; Patent No. 6060283
; GENERAL INFORMATION:
; APPLICANT: Takanori OKURA
; APPLICANT: Kakuji TORIGOE
; APPLICANT: Masahi KURIMOTO
; TITLE OF INVENTION: GENOMIC DNA ENCODING A POLYPEPTIDE CAPABLE
; TITLE OF INVENTION: OF INDUCING THE PRODUCTION OF INTERFERON-
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/884,324
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 185,305/96
; FILING DATE: 27-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: OKURA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4773 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: human
; TISSUE TYPE: placenta
; FEATURE:
; NAME/KEY: Intron
; LOCATION: 1..4773
; IDENTIFICATION METHOD: E
US-08-884-324-9

Query Match      2.0%; Score 34; DB 3; Length 4773;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 78 GTGGCTCATGCGCTATATATCCAGCACTTTGGGAG 111
|||||
Db 1979 GTGGCTCATGCGCTATATATCCAGCACTTTGGGAG 2012
```

```

RESULT 37
US-09-242-948-3/c
; Sequence 3, Application US/09242948
; Patent No. 6252057
; GENERAL INFORMATION:
; APPLICANT: Brady, Matthew J
; APPLICANT: Printen, John A
; APPLICANT: Saltiel, Alan R
; APPLICANT: Warner-Lambert Company,
; (Outside USA)
; TITLE OF INVENTION: Protein Targeting to Glycogen
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 201 Tabor Road
; CITY: Morris Plains
; STATE: NJ
; COUNTRY: US
; ZIP: 07950
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/242,948
; FILING DATE: 25-Feb-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/025,107
; FILING DATE: 30-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ashbrook, Charles W
; REFERENCE/DOCKET NUMBER: 5485-01-CA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 313 996-5215
; TELEFAX: 313 996-1553
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5789 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 4238..5176
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-242-948-3

Query Match      2.0%; Score 34; DB 4; Length 5789;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 153 AGCGTGGCCCAACATGGTGAACCCCTATCTCTACT 186
|||||
Db 401 AGCGTGGCCCAACATGGTGAACCCCTATCTCTACT 368

RESULT 38
US-08-884-324-13
; Sequence 13, Application US/08884324
; Patent No. 6060283
; GENERAL INFORMATION:
; APPLICANT: Takanori OKURA
; APPLICANT: Kakuji TORIGOE
; APPLICANT: Masahi KURIMOTO
; TITLE OF INVENTION: GENOMIC DNA ENCODING A POLYPEPTIDE CAPABLE
; TITLE OF INVENTION: OF INDUCING THE PRODUCTION OF INTERFERON-
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: BROWDY AND NEIMARK  
STREET: 419 Seventh Street, N.W., Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/884,324  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 185,305/96  
FILING DATE: 27-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: BROWDY, Roger L.  
REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: OKURA=1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11464 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORIGINAL SOURCE:  
ORGANISM: human  
TISSUE TYPE: placenta  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..3  
IDENTIFICATION METHOD: E  
NAME/KEY: leader peptide  
LOCATION: 4..82  
IDENTIFICATION METHOD: S  
NAME/KEY: intron  
LOCATION: 83..1453  
IDENTIFICATION METHOD: E  
NAME/KEY: leader peptide  
LOCATION: 1454..1465  
IDENTIFICATION METHOD: S  
NAME/KEY: intron  
LOCATION: 1466..4848  
IDENTIFICATION METHOD: E  
NAME/KEY: leader peptide  
LOCATION: 4849..4865  
IDENTIFICATION METHOD: S  
NAME/KEY: mat peptide  
LOCATION: 4866..4983  
IDENTIFICATION METHOD: S  
NAME/KEY: intron  
LOCATION: 4984..6317  
IDENTIFICATION METHOD: E  
NAME/KEY: mat peptide  
LOCATION: 6318..6451  
IDENTIFICATION METHOD: S  
NAME/KEY: intron  
LOCATION: 6452..11224  
IDENTIFICATION METHOD: E  
NAME/KEY: mat peptide  
LOCATION: 11225..11443  
IDENTIFICATION METHOD: S  
NAME/KEY: 3'UTR  
LOCATION: 11444..11464  
IDENTIFICATION METHOD: E

Query Match 2.0%; Score 34; DB 3; Length 11464;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 78 GTGGCTCATGCTATAATCCAGCAGCACTTTGGGAG 111  
|||||  
Db 8430 GTGGCTCATGCTATAATCCAGCAGCACTTTGGGAG 8463  
|||||  
RESULT 39  
US-09-087-465-3/c  
; Sequence 3, Application US/09087465A  
; Patent No. 6160092  
; GENERAL INFORMATION:  
; APPLICANT: Vinkemeier, Dwe  
; APPLICANT: Chen, Xiaomin  
; APPLICANT: Darnell Jr., James E  
; APPLICANT: Kuriyan, John  
; TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS OF  
; TITLE OF INVENTION: USE  
; FILE REFERENCE: 600-1-229  
; CURRENT APPLICATION NUMBER: US/09/087,465A  
; CURRENT FILING DATE: 1998-05-29  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 17949  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-087-465-3  
Query Match 2.0%; Score 34; DB 4; Length 17949;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 78 GTGGCTCATGCTATAATCCAGCAGCACTTTGGGAG 111  
|||||  
Db 17855 GTGGCTCATGCTATAATCCAGCAGCACTTTGGGAG 17822  
|||||  
RESULT 40  
US-08-884-324-14  
; Sequence 14, Application US/08884324  
; Patent No. 6060283  
; GENERAL INFORMATION:  
; APPLICANT: Takanori OKURA  
; APPLICANT: Kakuji TORIGOE  
; APPLICANT: Masahi KURIMOTO  
; TITLE OF INVENTION: GENOMIC DNA ENCODING A POLYPEPTIDE CAPABLE  
; TITLE OF INVENTION: OF INDUCING THE PRODUCTION OF INTERFERON-  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BROWDY AND NEIMARK  
; STREET: 419 Seventh Street, N.W., Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/884,324  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 185,305/96  
; FILING DATE: 27-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BROWDY, Roger L.



REGISTRATION NUMBER: 25,618  
REFERENCE/DOCKET NUMBER: OKURA-1  
TELEPHONE: 202-628-5197  
TELEFAX: 202-737-3528  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 28994 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
ORIGINAL SOURCE: human  
TISSUE TYPE: placenta  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..15606  
IDENTIFICATION METHOD: E  
NAME/KEY: leader peptide  
LOCATION: 15607..15685  
IDENTIFICATION METHOD: S  
NAME/KEY: intron  
LOCATION: 15686..17056  
IDENTIFICATION METHOD: E  
NAME/KEY: leader peptide  
LOCATION: 17057..17068  
IDENTIFICATION METHOD: S  
NAME/KEY: intron  
LOCATION: 17069..20451  
IDENTIFICATION METHOD: E  
NAME/KEY: leader peptide  
LOCATION: 20452..20468  
IDENTIFICATION METHOD: S  
NAME/KEY: mat peptide  
LOCATION: 20469..20586  
IDENTIFICATION METHOD: S  
NAME/KEY: intron  
LOCATION: 20587..21920  
IDENTIFICATION METHOD: E  
NAME/KEY: mat peptide  
LOCATION: 21921..22054  
IDENTIFICATION METHOD: S  
NAME/KEY: intron  
LOCATION: 22055..26827  
IDENTIFICATION METHOD: E  
NAME/KEY: mat peptide  
LOCATION: 26828..27046  
IDENTIFICATION METHOD: S  
NAME/KEY: 3'UTR  
LOCATION: 27047..28994  
IDENTIFICATION METHOD: E  
US-08-884-324-14

Query Match 2.0%; Score 34; DB 3; Length 28994;  
Best Local Similarity 100.0%; Pred. No. 1.7e-05;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GTGGCTCATGCTATAATCCAGCACTTTGGGAG 111  
Db 24033 GTGGCTATGCTATAATCCAGCACTTTGGGAG 24066

RESULT 41  
US-09-813-817-3  
Sequence 3, Application US/09813817  
Patent No. 6340583  
GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL001178

CURRENT APPLICATION NUMBER: US/09/813,817  
CURRENT FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 59065  
TYPE: DNA  
ORGANISM: Human  
US-09-813-817-3

Query Match 2.0%; Score 34; DB 4; Length 59065;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GTGGCTCATGCTATAATCCAGCACTTTGGGAG 111  
Db 13652 GTGGCTATGCTATAATCCAGCACTTTGGGAG 13685

RESULT 42  
US-09-978-197-3  
Sequence 3, Application US/09978197  
Patent No. 6403353  
GENERAL INFORMATION:  
APPLICANT: YAN, Chunhua et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
TITLE OF INVENTION: THEREOF  
FILE REFERENCE: CL001178DIV  
CURRENT APPLICATION NUMBER: US/09/978,197  
CURRENT FILING DATE: 2001-10-17  
PRIOR APPLICATION NUMBER: 09/813,817  
PRIOR FILING DATE: 2001-03-22  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 59065  
TYPE: DNA  
ORGANISM: Human  
US-09-978-197-3

Query Match 2.0%; Score 34; DB 4; Length 59065;  
Best Local Similarity 100.0%; Pred. No. 1.6e-05;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 GTGGCTCATGCTATAATCCAGCACTTTGGGAG 111  
Db 13652 GTGGCTATGCTATAATCCAGCACTTTGGGAG 13685

RESULT 43  
US-09-345-882-1  
Sequence 1, Application US/09345882  
Patent No. 6399373  
GENERAL INFORMATION:  
APPLICANT: Bougueleret, Lydie  
TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-  
AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.

FILE REFERENCE: GENSET.031A  
CURRENT APPLICATION NUMBER: US/09/345,882  
CURRENT FILING DATE: 1999-06-30  
PRIOR APPLICATION NUMBER: US 60/091,315  
PRIOR FILING DATE: 1998-06-30  
PRIOR APPLICATION NUMBER: US 60/111,909  
PRIOR FILING DATE: 1998-12-10  
NUMBER OF SEQ ID NOS: 140  
SOFTWARE: Patent.pm  
SEQ ID NO 1  
LENGTH: 162450  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: allele

```
LOCATION: 72794
OTHER INFORMATION: 5-124-273 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 88073
OTHER INFORMATION: 5-127-261 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 90842
OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97122
OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108308
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
FEATURE:
NAME/KEY: allele
LOCATION: 88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
```

```
;
; NAME/KEY: allele
; LOCATION: 103783..103828
; OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103783..103828
; OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106918..106966
; OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106918..106966
; OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108084..108130
; OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108084..108130
; OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108127..108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108127..108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
; FEATURE:
;
Query Match 2.0%; Score 34; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred. No. 1.5e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 GAGCTGGTGCACACACCTGTAATCCAGCTACT 237
|||||
Db 125653 GAGCTGGTGCACACACCTGTAATCCAGCTACT 125686

RESULT 44
US-08-991-789A-15/c
; Sequence 15, Application US/08991789A
; Patent No. 6225054
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
;
NUMBER OF SEQUENCES: 292
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed IP Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
;
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,789A
FILING DATE: 11-Dec-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Potter, Jane E. R.
REGISTRATION NUMBER: 33,332
```

```
;
; REFERENCE/DOCKET NUMBER: 210121.419C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-08-991-789A-15

Query Match 1.9%; Score 32; DB 4; Length 548;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 AGCACTTTGGAGCCTGAGGTGGGTGGATCAC 130
|||||
Db 483 AGCACTTTGGAGCCTGAGGTGGGTGGATCAC 452

RESULT 45
US-09-062-451-15/c
; Sequence 15, Application US/09062451
; Patent No. 6344550
; GENERAL INFORMATION:
; APPLICANT: Frudakis, Tony N.
; Smith, John M.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF BREAST CANCER
;
NUMBER OF SEQUENCES: 297
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
;
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/062,451
FILING DATE: 04-APR-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.419C2
;
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
;
INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 548 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-062-451-15

Query Match 1.9%; Score 32; DB 4; Length 548;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 AGCACTTTGGAGCCTGAGGTGGGTGGATCAC 130
|||||
Db 483 AGCACTTTGGAGCCTGAGGTGGGTGGATCAC 452
```

RESULT 46

US-09-598-326-15/c  
; Sequence 15, Application US/09598326  
; Patent No. 6423496  
; GENERAL INFORMATION:  
; APPLICANT: Fridakis, Tony N.  
; Smith, John M.  
; Reed, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF BREAST CANCER  
; NUMBER OF SEQUENCES: 247  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed Intellectual Property Law Group PLLC  
; STREET: 701 Fifth Avenue, Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/598,326  
; FILING DATE: 20-Jun-2000  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Potter, Jane E.R.  
; REGISTRATION NUMBER: 33,332  
; REFERENCE/DOCKET NUMBER: 210121.419D1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 548 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-09-598-326-15

Query Match 1.9%; Score 32; DB 4; Length 548;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 AGCACTTTGGAGCCTCAGGTGGTGATCAC 130  
|||||  
Db 483 AGCACTTTGGAGCCTCAGGTGGTGATCAC 452

RESULT 47

US-09-039-555B-19/c  
; Sequence 19, Application US/09039555B  
; Patent No. 6033856  
; GENERAL INFORMATION:  
; APPLICANT: Koerner, Kathrin  
; APPLICANT: Mueller, Rolf  
; APPLICANT: Sadlacek, Hans-Harald  
; TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,555B  
FILING DATE: 16-MAR-1998  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 19710643.9  
FILING DATE: 14-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 016779/0131  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2000 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-039-555B-19

Query Match 1.9%; Score 32; DB 3; Length 2000;  
Best Local Similarity 100.0%; Pred. No. 0.00014;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 GTGTGTGCACACACCTGTATCCAGCTACTT 238  
|||||  
Db 913 GTGTGTGCACACACCTGTATCCAGCTACTT 882

RESULT 48

US-09-851-896-3/c  
; Sequence 3, Application US/09851896  
; Patent No. 6410325  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Susan M. Freier  
; APPLICANT: Andrew T. Watt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEP  
; TITLE OF INVENTION: EXPRESSION  
; FILE REFERENCE: ITS-0220  
; CURRENT APPLICATION NUMBER: US/09/851,896  
; CURRENT FILING DATE: 2001-05-08  
; NUMBER OF SEQ ID NOS: 89  
; SEQ ID NO 3  
; LENGTH: 70000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-851-896-3

Query Match 1.9%; Score 32; DB 4; Length 70000;  
Best Local Similarity 100.0%; Pred. No. 0.00011;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GGCTCATGCTATATCCAGCAGCTTTGGGAG 111  
|||||  
Db 29902 GGCTCATGCTATATCCAGCAGCTTTGGGAG 29871

RESULT 49

US-09-128-155-17  
; Sequence 17, Application US/09128155  
; Patent No. 6117654  
; GENERAL INFORMATION:  
; APPLICANT: Pan, Yang  
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: 09404/052001  
; CURRENT APPLICATION NUMBER: US/09/128,155  
; CURRENT FILING DATE: 1998-08-03  
; EARLIER APPLICATION NUMBER: US 60/091,650  
; EARLIER FILING DATE: 1998-07-02  
; EARLIER APPLICATION NUMBER: US 60/054,646  
; EARLIER FILING DATE: 1997-08-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 17  
; LENGTH: 176373  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(176373)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-17

Query Match 1.9%; Score 32; DB 3; Length 176373;  
Best Local Similarity 100.0%; Pred. No. 0.0001;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 GGCTCATGCTATAATCCAGCAGCTTTGGGAG 111  
Db 127041 GGCTCATGCTATAATCCAGCAGCTTTGGGAG 127072

RESULT 50  
US-09-183-266A-12/C  
; Sequence 12, Application US/09183266A  
; Patent No. 6361954  
; GENERAL INFORMATION:  
; APPLICANT: Stillman, Bruce  
; APPLICANT: Williams, R. Sanders  
; APPLICANT: Mendez, Juan  
; TITLE OF INVENTION: DNA REPLICATION-REGULATING GENES,  
; TITLE OF INVENTION: ANTIBODIES THERETO AND DIAGNOSTIC APPLICATIONS THEREOF  
; FILE REFERENCE: CSHL96-01A3  
; CURRENT APPLICATION NUMBER: US/09/183,266A  
; CURRENT FILING DATE: 1998-10-30  
; PRIOR APPLICATION NUMBER: PCT/US97/07333  
; PRIOR FILING DATE: 1997-05-02  
; PRIOR APPLICATION NUMBER: 08/648,650  
; PRIOR FILING DATE: 1996-05-15  
; PRIOR APPLICATION NUMBER: 08/643,034  
; PRIOR FILING DATE: 1996-05-02  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 1210  
; TYPE: DNA  
; ORGANISM: H. sapiens  
US-09-183-266A-12

Query Match 1.8%; Score 31; DB 4; Length 1210;  
Best Local Similarity 100.0%; Pred. No. 0.00039;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GCTCATGCTATAATCCAGCAGCTTTGGGAG 111  
Db 160 GCTCATGCTATAATCCAGCAGCTTTGGGAG 130

Search completed: February 12, 2003, 16:27:34  
Job time : 2065 secs



GenCore version 5.1.3  
 Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2003, 08:12:30 ; Search time 87 Seconds  
 (without alignments)  
 9835.215 Million cell updates/sec

Title: US-09-942-310-2  
 Perfect score: 1680  
 Sequence: 1 gaattcaagaccagctgga.....catctctctgctctggtgg 1680

Scoring table: OLIGO\_NUC  
 Gapop 60.0 , Gapext 60.0

Searched: 424239 seqs, 254661826 residues

Word size : 20

Total number of hits satisfying chosen parameters: 22880

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/UCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/UCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length	ID	Description
C 1	49	2.9	32177	10 US-09-764-877-3251	Sequence 3251, Ap
C 2	48	2.9	65464	9 US-09-859-888-3	Sequence 3, Appli
C 3	47	2.8	776	10 US-09-728-711-7	Sequence 7, Appli
C 4	46	2.7	133893	9 US-10-161-510-1	Sequence 1, Appli
C 5	43	2.6	58837	10 US-09-982-091A-5	Sequence 5, Appli
C 6	42	2.5	397	10 US-09-867-701-8121	Sequence 8121, Ap
C 7	42	2.5	302250	10 US-09-962-832-154	Sequence 154, App
C 8	41	2.4	400	10 US-09-867-701-7324	Sequence 7324, Ap
C 9	41	2.4	2694	10 US-09-880-107-3872	Sequence 3872, Ap
C 10	41	2.4	3088	10 US-09-954-456-45	Sequence 45, Appli
C 11	41	2.4	3088	10 US-09-954-456-1621	Sequence 1621, Ap
C 12	41	2.4	3088	10 US-09-969-347-234	Sequence 234, App
C 13	41	2.4	15300	10 US-09-764-860-1091	Sequence 1091, Ap
C 14	41	2.4	16552	10 US-09-764-855-321	Sequence 321, App
C 15	41	2.4	16552	10 US-09-764-855-332	Sequence 332, App
C 16	41	2.4	18878	10 US-09-764-877-3806	Sequence 3806, App
C 17	41	2.4	57130	10 US-09-835-081-3	Sequence 3, Appli
C 18	41	2.4	58837	10 US-09-982-091A-5	Sequence 5, Appli
C 19	41	2.4	58985	9 US-09-901-152-3	Sequence 3, Appli

2.4	65464	9	US-09-859-888-3	Sequence 3, Appli
2.4	76798	10	US-09-880-107-3949	Sequence 3949, Ap
2.4	111282	12	US-10-094-989-3	Sequence 3, Appli
2.4	126512	10	US-09-804-474A-3	Sequence 3, Appli
2.4	145831	10	US-09-969-708-79	Sequence 79, Appli
2.4	145831	10	US-09-954-456-2116	Sequence 2116, Ap
2.4	470	10	US-09-880-107-832	Sequence 832, App
2.4	2791	10	US-09-729-674-51	Sequence 51, Appli
2.4	5815	10	US-09-764-847-1487	Sequence 1487, Ap
2.4	7373	10	US-09-764-853-896	Sequence 896, App
2.4	9958	10	US-09-764-877-2718	Sequence 2718, Ap
2.4	10342	9	US-09-764-868-1471	Sequence 1471, Ap
2.4	13069	10	US-09-764-869-1850	Sequence 1850, Ap
2.4	27377	10	US-09-816-248-18	Sequence 18, Appli
2.4	28770	10	US-09-817-198A-3	Sequence 3, Appli
2.4	32203	10	US-09-764-869-1849	Sequence 1849, Ap
2.4	51719	10	US-09-918-686-2	Sequence 2, Appli
2.4	84539	10	US-09-962-436-36	Sequence 36, Appli
2.4	92139	10	US-09-918-686-1	Sequence 1, Appli
2.4	110096	10	US-09-880-107-1542	Sequence 1542, Ap
2.3	349	10	US-09-764-877-2735	Sequence 2735, Ap
2.3	425	10	US-09-764-877-2736	Sequence 2736, Ap
2.3	1788	10	US-09-822-830A-569	Sequence 569, App
2.3	8082	10	US-09-764-860-1106	Sequence 1106, Ap
2.3	32190	10	US-09-764-869-2209	Sequence 2209, Ap
2.3	26048	10	US-09-764-869-1556	Sequence 1556, Ap
2.3	143068	10	US-09-967-768A-316	Sequence 316, App
2.3	167343	10	US-09-962-436-281	Sequence 281, App
2.3	167343	10	US-09-964-824A-273	Sequence 273, App
2.2	293	10	US-09-764-877-376	Sequence 376, App
2.2	438	10	US-09-867-701-5962	Sequence 5962, Ap
2.2	643	10	US-09-764-847-1024	Sequence 1024, Ap
2.2	1174	10	US-09-764-877-2611	Sequence 2611, Ap
2.2	8220	10	US-09-797-908-3	Sequence 3, Appli
2.2	11469	10	US-09-764-877-2791	Sequence 2791, Ap
2.2	22484	10	US-09-875-114-2	Sequence 2, Appli
2.2	22484	10	US-09-880-107-3341	Sequence 3341, Ap
2.2	26657	10	US-09-810-673A-3	Sequence 3, Appli
2.2	174424	10	US-09-967-768A-314	Sequence 314, App
2.1	323	10	US-09-867-701-9244	Sequence 9244, Ap
2.1	384	10	US-09-867-701-7719	Sequence 7719, Ap
2.1	471	10	US-09-867-701-819	Sequence 819, App
2.1	1282	10	US-09-822-830A-319	Sequence 319, App
2.1	4457	10	US-09-764-877-3682	Sequence 3682, Ap
2.1	19616	10	US-09-764-877-3220	Sequence 3220, Ap
2.1	20532	10	US-09-764-877-3774	Sequence 3774, Ap
2.1	23071	10	US-09-764-864-1673	Sequence 1673, Ap
2.1	28818	10	US-09-764-877-2266	Sequence 2266, Ap
2.1	30013	10	US-09-764-877-3297	Sequence 3297, Ap
2.1	49744	10	US-09-927-091-4	Sequence 4, Appli
2.1	52354	10	US-09-742-311-3	Sequence 3, Appli
2.1	99014	10	US-09-880-107-3428	Sequence 3428, Ap
2.1	106323	10	US-09-803-661-3	Sequence 3, Appli
2.1	152331	9	US-10-095-407-16	Sequence 16, Appli
2.1	202001	10	US-09-734-674-3	Sequence 3, Appli
2.1	3009	10	US-09-867-701-9236	Sequence 9236, Ap
2.1	312	10	US-09-867-701-10749	Sequence 10749, A
2.1	321	10	US-09-867-701-10735	Sequence 10735, A
2.1	339	10	US-09-867-701-9343	Sequence 9343, Ap
2.1	378	10	US-09-867-701-7421	Sequence 7421, Ap
2.1	386	10	US-09-969-347-293	Sequence 293, App
2.1	386	10	US-09-867-701-8177	Sequence 8177, Ap
2.1	397	10	US-09-920-300A-1579	Sequence 1579, Ap
2.1	397	12	US-10-033-528-1579	Sequence 1579, Ap
2.1	451	10	US-09-867-701-10468	Sequence 10468, A
2.1	471	10	US-09-867-701-6782	Sequence 6782, Ap
2.1	903	10	US-09-764-877-2956	Sequence 2956, Ap
2.1	1779	10	US-09-764-860-774	Sequence 774, App
2.1	1780	10	US-09-764-860-773	Sequence 773, App
2.1	1988	10	US-09-764-869-1788	Sequence 1788, Ap
2.1	6040	10	US-09-764-869-1029	Sequence 1029, Ap
2.1	6235	9	US-09-845-020A-5	Sequence 5, Appli
2.1	6427	10	US-09-764-855-292	Sequence 292, App

93	35	2.1	6544	10	US-09-764-847-1544	Sequence 1544, Ap	166	34	2.0	3582	9	US-10-176-747-465	Sequence 465, App
94	35	2.1	6679	9	US-09-845-020A-1	Sequence 1, Appli	167	34	2.0	3582	9	US-10-176-750-465	Sequence 465, App
c 95	35	2.1	6892	9	US-09-764-877-3770	Sequence 3770, Ap	168	34	2.0	3582	9	US-10-176-985-465	Sequence 465, App
96	35	2.1	13274	10	US-09-764-877-2428	Sequence 2428, Ap	169	34	2.0	3582	9	US-10-176-987-465	Sequence 465, App
c 97	35	2.1	13953	9	US-10-096-961-3	Sequence 3, Appli	170	34	2.0	3582	9	US-10-176-991-465	Sequence 465, App
c 98	35	2.1	16106	10	US-09-764-877-2322	Sequence 2322, Ap	171	34	2.0	3582	9	US-10-176-992-465	Sequence 465, App
c 99	35	2.1	18466	10	US-09-764-869-1682	Sequence 1682, Ap	172	34	2.0	3582	9	US-10-176-993-465	Sequence 465, App
c 100	35	2.1	19846	10	US-09-764-869-1683	Sequence 1683, Ap	173	34	2.0	3582	9	US-10-184-658-465	Sequence 465, App
c 101	35	2.1	23934	10	US-09-764-877-2536	Sequence 2536, Ap	174	34	2.0	3582	12	US-10-053-586-465	Sequence 465, App
c 102	35	2.1	23934	10	US-09-764-869-1683	Sequence 2544, Ap	c 175	34	2.0	4963	10	US-09-764-870-606	Sequence 606, App
c 103	35	2.1	23934	10	US-09-764-877-2536	Sequence 3, Appli	176	34	2.0	4963	10	US-09-764-860-694	Sequence 694, App
c 104	35	2.1	24707	10	US-09-764-877-2544	Sequence 777, App	177	34	2.0	6107	10	US-09-764-869-2088	Sequence 2088, Ap
c 105	35	2.1	25629	12	US-10-135-689-3	Sequence 775, App	178	34	2.0	7347	10	US-09-764-869-2089	Sequence 2088, Ap
c 106	35	2.1	31885	10	US-09-764-860-775	Sequence 2530, Ap	179	34	2.0	7351	10	US-09-764-869-2090	Sequence 2090, Ap
c 107	35	2.1	31885	10	US-09-764-877-2530	Sequence 3, Appli	180	34	2.0	7626	9	US-10-001-835-82	Sequence 82, Appli
c 108	35	2.1	31885	10	US-09-764-877-2541	Sequence 2541, Ap	c 181	34	2.0	8746	10	US-09-764-860-1022	Sequence 1022, Ap
c 109	35	2.1	32207	10	US-09-764-877-3250	Sequence 3250, Ap	c 182	34	2.0	10739	10	US-09-764-869-2130	Sequence 2130, Ap
c 110	35	2.1	43950	12	US-10-060-332-3	Sequence 3, Appli	c 183	34	2.0	11185	10	US-09-764-860-1096	Sequence 1096, Ap
c 111	35	2.1	48436	10	US-09-927-602-38	Sequence 38, Appli	184	34	2.0	11464	12	US-10-100-057-17	Sequence 17, Appli
c 112	35	2.1	63588	9	US-10-243-735-3	Sequence 3, Appli	c 185	34	2.0	11990	10	US-09-963-708-569	Sequence 569, App
c 113	35	2.1	63588	9	US-10-243-735-3	Sequence 3, Appli	c 186	34	2.0	15366	10	US-09-764-877-3797	Sequence 3797, Ap
c 114	35	2.1	65608	10	US-09-954-531-180	Sequence 180, App	c 187	34	2.0	15371	10	US-09-764-877-3798	Sequence 3798, Ap
c 115	35	2.1	65608	10	US-09-962-436-292	Sequence 292, App	c 188	34	2.0	16100	10	US-09-764-877-3698	Sequence 3698, Ap
c 116	35	2.1	65608	10	US-09-962-436-292	Sequence 119, App	c 189	34	2.0	17216	10	US-09-764-877-3565	Sequence 3565, Ap
c 117	35	2.1	81001	10	US-09-751-877-1	Sequence 1, Appli	c 190	34	2.0	17217	10	US-09-764-877-3566	Sequence 3566, Ap
c 118	35	2.1	90541	10	US-09-759-359A-3	Sequence 29, Appli	c 191	34	2.0	18648	10	US-09-954-456-1150	Sequence 1150, Ap
c 119	34	2.0	214	10	US-09-867-701-9610	Sequence 2202, Ap	c 192	34	2.0	22452	9	US-09-764-868-1487	Sequence 1487, Ap
c 120	34	2.0	260	9	US-09-736-457-1208	Sequence 2259, Ap	c 193	34	2.0	22452	9	US-09-764-868-1489	Sequence 1489, Ap
c 121	34	2.0	260	9	US-09-902-941-1208	Sequence 1208, Ap	c 194	34	2.0	23822	10	US-09-964-824A-572	Sequence 572, App
c 122	34	2.0	260	9	US-09-849-626-1208	Sequence 1208, Ap	c 195	34	2.0	27681	10	US-09-764-869-1997	Sequence 1997, Ap
c 123	34	2.0	293	10	US-09-764-869-1208	Sequence 498, App	c 196	34	2.0	27681	10	US-09-764-869-1998	Sequence 1998, Ap
c 124	34	2.0	356	10	US-09-867-701-3326	Sequence 3326, App	c 197	34	2.0	28588	10	US-09-764-887-399	Sequence 399, App
c 125	34	2.0	378	10	US-09-867-701-862	Sequence 862, App	c 198	34	2.0	31348	10	US-09-764-869-1259	Sequence 1259, Ap
c 126	34	2.0	490	10	US-09-783-590-1069	Sequence 1069, Ap	c 199	34	2.0	31394	9	US-09-764-904-71	Sequence 71, Appli
c 127	34	2.0	998	9	US-10-016-634A-29	Sequence 29, Appli	c 200	34	2.0	31994	10	US-09-764-860-599	Sequence 599, App
c 128	34	2.0	1022	10	US-09-764-869-2202	Sequence 2202, Ap	c 201	34	2.0	32152	10	US-09-764-855-328	Sequence 328, App
c 129	34	2.0	1022	10	US-09-764-869-2202	Sequence 2259, Ap	c 202	34	2.0	32204	10	US-09-764-855-327	Sequence 327, App
c 130	34	2.0	1042	10	US-09-764-869-2269	Sequence 1420, Ap	c 203	34	2.0	45862	9	US-10-216-355-3	Sequence 3, Appli
c 131	34	2.0	1268	10	US-09-764-847-1420	Sequence 3388, Ap	c 204	34	2.0	53542	10	US-09-801-574-61	Sequence 61, Appli
c 132	34	2.0	1268	10	US-09-764-877-3388	Sequence 3388, Ap	c 205	34	2.0	56377	10	US-09-782-378A-17	Sequence 17, Appli
c 133	34	2.0	1268	10	US-09-764-877-3388	Sequence 3389, Ap	c 206	34	2.0	57467	9	US-10-237-859-3	Sequence 3, Appli
c 134	34	2.0	1529	10	US-09-764-860-897	Sequence 897, App	c 207	34	2.0	88191	10	US-09-799-799-3	Sequence 3, Appli
c 135	34	2.0	1529	10	US-09-764-860-898	Sequence 898, App	c 208	34	2.0	92139	10	US-09-918-686-1	Sequence 1, Appli
c 136	34	2.0	1529	10	US-09-764-860-899	Sequence 899, App	c 209	34	2.0	143068	10	US-09-967-768A-316	Sequence 316, App
c 137	34	2.0	1529	10	US-09-764-860-900	Sequence 900, App	c 210	34	2.0	174424	10	US-09-967-768A-314	Sequence 314, App
c 138	34	2.0	2571	10	US-09-764-877-2215	Sequence 2215, Ap	c 211	34	2.0	203654	10	US-09-820-905-3	Sequence 3, Appli
c 139	34	2.0	2571	10	US-09-764-877-2216	Sequence 2216, Ap	c 212	34	2.0	326014	10	US-09-731-231A-3	Sequence 3, Appli
c 140	34	2.0	2834	9	US-09-764-877-2217	Sequence 2217, Ap	c 213	34	2.0	465237	10	US-09-933-267A-1	Sequence 1, Appli
c 141	34	2.0	3498	10	US-09-845-020A-6	Sequence 6, Appli	c 214	34	2.0	684973	10	US-09-263-959-1	Sequence 1, Appli
c 142	34	2.0	3498	10	US-09-764-869-2204	Sequence 2204, Ap	c 215	34	2.0	1503841	9	US-09-946-807-1	Sequence 1, Appli
c 143	34	2.0	3498	10	US-09-764-869-2271	Sequence 2271, Ap	c 216	34	2.0	1503841	10	US-09-795-668-1	Sequence 1, Appli
c 144	34	2.0	3499	10	US-09-764-869-2203	Sequence 2270, Ap	c 217	34	2.0	1503841	10	US-09-795-668-1	Sequence 1, Appli
c 145	34	2.0	3582	9	US-10-174-590-465	Sequence 465, App	c 218	33	2.0	233	10	US-09-867-701-10133	Sequence 10133, A
c 146	34	2.0	3582	9	US-10-176-758-465	Sequence 465, App	c 219	33	2.0	349	10	US-09-867-701-8507	Sequence 8507, Ap
c 147	34	2.0	3582	9	US-10-175-737-465	Sequence 465, App	c 220	33	2.0	376	10	US-09-908-711-137	Sequence 137, App
c 148	34	2.0	3582	9	US-10-173-706-465	Sequence 465, App	c 221	33	2.0	376	10	US-09-908-711-138	Sequence 138, App
c 149	34	2.0	3582	9	US-10-175-738-465	Sequence 465, App	c 222	33	2.0	401	9	US-09-946-807-917	Sequence 917, App
c 150	34	2.0	3582	9	US-10-175-738-465	Sequence 465, App	c 223	33	2.0	401	10	US-09-795-668-917	Sequence 917, App
c 151	34	2.0	3582	9	US-10-175-732-465	Sequence 465, App	c 224	33	2.0	401	10	US-09-795-668-917	Sequence 917, App
c 152	34	2.0	3582	9	US-10-176-482-465	Sequence 465, App	c 225	33	2.0	495	10	US-09-867-701-703	Sequence 703, App
c 153	34	2.0	3582	9	US-10-176-757-465	Sequence 465, App	c 226	33	2.0	498	9	US-09-292-758-138	Sequence 138, App
c 154	34	2.0	3582	9	US-10-176-913-465	Sequence 465, App	c 227	33	2.0	573	10	US-09-764-869-1483	Sequence 1483, Ap
c 155	34	2.0	3582	9	US-10-180-552-465	Sequence 465, App	c 228	33	2.0	573	10	US-09-764-869-1484	Sequence 1484, Ap
c 156	34	2.0	3582	9	US-10-180-557-465	Sequence 465, App	c 229	33	2.0	2264	10	US-09-764-877-3364	Sequence 3364, Ap
c 157	34	2.0	3582	9	US-10-173-700-465	Sequence 465, App	c 230	33	2.0	5744	10	US-09-764-847-1457	Sequence 1457, Ap
c 158	34	2.0	3582	9	US-10-174-572-465	Sequence 465, App	c 231	33	2.0	7233	10	US-09-764-869-2123	Sequence 2123, Ap
c 159	34	2.0	3582	9	US-10-174-579-465	Sequence 465, App	c 232	33	2.0	7233	10	US-09-764-869-2123	Sequence 2123, Ap
c 160	34	2.0	3582	9	US-10-174-582-465	Sequence 465, App	c 233	33	2.0	9692	9	US-09-764-869-1399	Sequence 1399, Ap
c 161	34	2.0	3582	9	US-10-174-588-465	Sequence 465, App	c 234	33	2.0	9730	9	US-09-764-868-1400	Sequence 1400, Ap
c 162	34	2.0	3582	9	US-10-175-739-465	Sequence 465, App	c 235	33	2.0	9883	9	US-10-016-157A-71	Sequence 71, Appli
c 163	34	2.0	3582	9	US-10-175-739-465	Sequence 465, App	c 236	33	2.0	10953	10	US-09-764-846-313	Sequence 313, App
c 164	34	2.0	3582	9	US-10-175-740-465	Sequence 465, App	c 237	33	2.0	11869	10	US-09-764-869-2292	Sequence 2292, Ap
c 165	34	2.0	3582	9	US-10-176-488-465	Sequence 465, App	c 238	33	2.0	14962	10	US-09-764-878-244	Sequence 244, App



c 239	33	2.0	16747	10	US-09-764-877-3354	Sequence 3354, Ap	c 312	31	1.8	168	9	US-09-860-670-229	Sequence 229, App
c 240	33	2.0	17752	10	US-09-748-127-3	Sequence 3, Appli	c 313	31	1.8	313	10	US-09-867-701-142	Sequence 142, App
c 241	33	2.0	18535	10	US-09-764-878-385	Sequence 385, App	c 314	31	1.8	324	10	US-09-867-701-7961	Sequence 7961, Ap
c 242	33	2.0	20272	10	US-09-908-711-145	Sequence 145, App	c 315	31	1.8	327	10	US-09-962-832-72	Sequence 72, Appl
c 243	33	2.0	21636	10	US-09-416-384A-3	Sequence 3, Appli	c 316	31	1.8	327	10	US-09-880-107-652	Sequence 652, App
c 244	33	2.0	23378	10	US-09-764-847-1514	Sequence 1514, Ap	c 317	31	1.8	336	10	US-09-867-701-7075	Sequence 7075, Ap
c 245	33	2.0	25619	10	US-09-908-711-143	Sequence 143, App	c 318	31	1.8	449	10	US-09-867-701-8942	Sequence 8942, Ap
c 246	33	2.0	25619	10	US-09-764-898-302	Sequence 302, App	c 319	31	1.8	1779	10	US-09-822-849A-133	Sequence 133, App
c 247	33	2.0	31814	10	US-09-817-182-3	Sequence 3, Appli	c 320	31	1.8	2105	10	US-09-939-825-15	Sequence 15, Appl
c 248	33	2.0	32154	10	US-09-764-877-3433	Sequence 3433, Ap	c 321	31	1.8	2756	10	US-09-764-853-998	Sequence 858, App
c 249	33	2.0	32170	10	US-09-764-860-1108	Sequence 1108, Ap	c 322	31	1.8	3064	10	US-09-764-847-1289	Sequence 1289, Ap
c 250	33	2.0	52216	10	US-09-747-810-1	Sequence 1, Appli	c 323	31	1.8	3064	10	US-09-764-847-1290	Sequence 1290, Ap
c 251	33	2.0	90541	10	US-09-759-359A-3	Sequence 3, Appli	c 324	31	1.8	3329	10	US-09-090-672B-5	Sequence 5, Appli
c 252	33	2.0	133893	9	US-10-161-510-1	Sequence 1, Appli	c 325	31	1.8	9365	12	US-10-092-063-8	Sequence 8, Appli
c 253	32	1.9	335	10	US-09-867-701-8670	Sequence 8670, Ap	c 326	31	1.8	9365	12	US-10-091-085-8	Sequence 8, Appli
c 254	32	1.9	340	10	US-09-867-701-10062	Sequence 10062, A	c 327	31	1.8	10563	10	US-09-764-864-1680	Sequence 1680, Ap
c 255	32	1.9	356	10	US-09-764-869-442	Sequence 442, App	c 328	31	1.8	11474	10	US-09-764-847-1559	Sequence 1559, Ap
c 256	32	1.9	382	10	US-09-867-701-10167	Sequence 10167, A	c 329	31	1.8	12477	10	US-09-764-870-559	Sequence 559, App
c 257	32	1.9	401	9	US-09-946-807-708	Sequence 708, App	c 330	31	1.8	12477	10	US-09-764-853-851	Sequence 851, App
c 258	32	1.9	401	10	US-09-795-668-708	Sequence 708, App	c 331	31	1.8	13953	9	US-10-096-961-3	Sequence 3, Appli
c 259	32	1.9	494	10	US-09-795-686-708	Sequence 708, App	c 332	31	1.8	18501	10	US-09-764-847-1916	Sequence 1916, Ap
c 260	32	1.9	494	10	US-09-867-701-6998	Sequence 6998, Ap	c 333	31	1.8	18860	10	US-09-764-877-2317	Sequence 2317, Ap
c 261	32	1.9	526	10	US-09-867-701-7000	Sequence 7000, Ap	c 334	31	1.8	18966	10	US-09-764-877-3372	Sequence 3372, Ap
c 262	32	1.9	548	9	US-09-924-400-15	Sequence 15, Appl	c 335	31	1.8	19820	10	US-09-764-877-2713	Sequence 2713, Ap
c 263	32	1.9	548	10	US-09-810-936-15	Sequence 15, Appl	c 336	31	1.8	20210	10	US-09-764-870-598	Sequence 598, App
c 264	32	1.9	548	10	US-09-429-755-15	Sequence 15, Appl	c 337	31	1.8	20268	10	US-09-764-855-173	Sequence 173, App
c 265	32	1.9	687	10	US-09-764-878-373	Sequence 373, App	c 338	31	1.8	23603	9	US-09-860-670-264	Sequence 264, App
c 266	32	1.9	1003	10	US-09-764-855-308	Sequence 308, App	c 339	31	1.8	23613	9	US-09-860-670-258	Sequence 258, App
c 267	32	1.9	1003	10	US-09-764-855-309	Sequence 309, App	c 340	31	1.8	24218	9	US-09-860-670-263	Sequence 263, App
c 268	32	1.9	1331	10	US-09-822-849A-113	Sequence 113, App	c 341	31	1.8	29607	10	US-09-764-877-3626	Sequence 3626, Ap
c 269	32	1.9	2012	10	US-09-764-877-3879	Sequence 3879, Ap	c 342	31	1.8	32169	10	US-09-764-847-1963	Sequence 1963, App
c 270	32	1.9	2012	10	US-09-764-877-3880	Sequence 3880, Ap	c 343	31	1.8	32188	10	US-09-764-860-799	Sequence 799, App
c 271	32	1.9	3341	9	US-09-764-868-1433	Sequence 1433, Ap	c 344	31	1.8	32190	9	US-09-860-670-255	Sequence 255, App
c 272	32	1.9	3342	9	US-09-764-868-1434	Sequence 1434, Ap	c 345	31	1.8	32248	10	US-09-764-864-1769	Sequence 1769, Ap
c 273	32	1.9	4645	10	US-09-764-869-1482	Sequence 1482, Ap	c 346	31	1.8	32248	10	US-09-764-877-3487	Sequence 3487, Ap
c 274	32	1.9	4646	10	US-09-764-869-1481	Sequence 1481, Ap	c 347	31	1.8	32249	9	US-09-860-670-260	Sequence 260, App
c 275	32	1.9	5000	10	US-09-791-105-1	Sequence 1, Appli	c 348	31	1.8	32816	10	US-09-729-094-3	Sequence 3, Appli
c 276	32	1.9	5251	10	US-09-764-869-1431	Sequence 1431, Ap	c 349	31	1.8	35414	10	US-09-836-607-45	Sequence 45, Appl
c 277	32	1.9	5363	10	US-09-764-878-229	Sequence 229, App	c 350	31	1.8	55795	10	US-09-880-107-1543	Sequence 1543, Ap
c 278	32	1.9	8868	10	US-09-764-877-3954	Sequence 3954, Ap	c 351	31	1.8	63000	10	US-09-780-172-18	Sequence 18, Appl
c 279	32	1.9	10476	10	US-09-964-824A-98	Sequence 98, Appl	c 352	31	1.8	66686	10	US-09-736-960-86	Sequence 86, Appl
c 280	32	1.9	10476	10	US-09-964-824A-552	Sequence 552, App	c 353	31	1.8	73467	9	US-10-237-859-3	Sequence 3, Appli
c 281	32	1.9	10867	10	US-09-764-869-2113	Sequence 2113, Ap	c 354	31	1.8	75270	9	US-09-790-852-1	Sequence 1, Appli
c 282	32	1.9	10867	10	US-09-764-877-3970	Sequence 3970, Ap	c 355	31	1.8	170830	10	US-09-835-232-7	Sequence 7, Appli
c 283	32	1.9	10907	10	US-09-764-869-2112	Sequence 2112, Ap	c 356	31	1.8	174493	9	US-10-238-709-3	Sequence 3, Appli
c 284	32	1.9	10907	10	US-09-764-877-3968	Sequence 3968, Ap	c 357	31	1.8	174493	9	US-09-804-471A-3	Sequence 3, Appli
c 285	32	1.9	11881	9	US-09-764-868-1351	Sequence 1351, Ap	c 358	30	1.8	83	10	US-09-764-877-3299	Sequence 3299, Ap
c 286	32	1.9	11881	9	US-09-764-868-1353	Sequence 1353, Ap	c 359	30	1.8	185	10	US-09-764-887-400	Sequence 400, App
c 287	32	1.9	11881	10	US-09-764-869-2124	Sequence 2124, Ap	c 360	30	1.8	185	10	US-09-764-887-401	Sequence 401, App
c 288	32	1.9	12718	10	US-09-764-877-3972	Sequence 3972, Ap	c 361	30	1.8	186	10	US-09-764-887-402	Sequence 402, App
c 289	32	1.9	17397	10	US-09-764-869-1945	Sequence 1945, Ap	c 362	30	1.8	191	10	US-09-764-847-997	Sequence 997, App
c 290	32	1.9	17397	10	US-09-764-869-1599	Sequence 1599, Ap	c 363	30	1.8	229	9	US-09-860-670-172	Sequence 172, App
c 291	32	1.9	19334	10	US-09-764-869-1943	Sequence 1943, Ap	c 364	30	1.8	303	10	US-09-764-860-776	Sequence 776, App
c 292	32	1.9	19345	10	US-09-764-869-1944	Sequence 1944, Ap	c 365	30	1.8	303	10	US-09-764-877-2535	Sequence 2535, Ap
c 293	32	1.9	19472	10	US-09-764-864-1698	Sequence 1698, Ap	c 366	30	1.8	303	10	US-09-764-877-2543	Sequence 2543, Ap
c 294	32	1.9	22161	10	US-09-764-847-1020	Sequence 1020, Ap	c 367	30	1.8	335	9	US-10-040-733-1433	Sequence 1433, Ap
c 295	32	1.9	29449	9	US-09-989-442-161	Sequence 161, App	c 368	30	1.8	366	9	US-09-796-692-8332	Sequence 8332, Ap
c 296	32	1.9	30175	10	US-09-738-878-3	Sequence 3, Appli	c 369	30	1.8	371	10	US-09-867-701-7806	Sequence 7806, Ap
c 297	32	1.9	30175	12	US-10-163-381-3	Sequence 3, Appli	c 370	30	1.8	378	9	US-09-768-827-74	Sequence 74, Appl
c 298	32	1.9	30350	9	US-10-118-328-3	Sequence 3, Appli	c 371	30	1.8	378	10	US-09-867-701-7743	Sequence 7743, Ap
c 299	32	1.9	32146	10	US-09-764-860-797	Sequence 797, App	c 372	30	1.8	381	10	US-09-867-701-10426	Sequence 10426, A
c 300	32	1.9	32150	10	US-09-764-877-3433	Sequence 3433, Ap	c 373	30	1.8	401	9	US-09-946-807-372	Sequence 372, App
c 301	32	1.9	32174	10	US-09-764-860-1108	Sequence 1108, Ap	c 374	30	1.8	401	10	US-09-795-668-372	Sequence 372, App
c 302	32	1.9	32195	10	US-09-764-869-2017	Sequence 2017, Ap	c 375	30	1.8	401	10	US-09-795-686-372	Sequence 372, App
c 303	32	1.9	32219	10	US-09-764-869-2016	Sequence 2016, Ap	c 376	30	1.8	420	9	US-10-040-739-12	Sequence 12, Appl
c 304	32	1.9	32221	10	US-09-764-847-1406	Sequence 1406, Ap	c 377	30	1.8	451	10	US-09-764-869-1764	Sequence 1764, Ap
c 305	32	1.9	32248	10	US-09-764-860-802	Sequence 802, App	c 378	30	1.8	451	10	US-09-764-860-789	Sequence 789, App
c 306	32	1.9	51552	10	US-09-733-294A-30	Sequence 30, Appl	c 379	30	1.8	467	9	US-10-202-193-122	Sequence 122, App
c 307	32	1.9	75899	10	US-09-854-883-243	Sequence 243, App	c 380	30	1.8	478	9	US-09-925-299-22	Sequence 22, Appl
c 308	32	1.9	84539	10	US-09-962-436-36	Sequence 36, Appl	c 381	30	1.8	495	9	US-10-202-193-215	Sequence 215, App
c 309	32	1.9	176373	9	US-10-095-407-17	Sequence 17, Appl	c 382	30	1.8	503	9	US-10-202-193-190	Sequence 190, App
c 310	32	1.9	203654	10	US-09-820-905-3	Sequence 3, Appli	c 383	30	1.8	534	10	US-09-764-877-3451	Sequence 3451, Ap
c 311	32	1.9	684973	10	US-09-263-959-1	Sequence 1, Appli	c 384	30	1.8	554	10	US-09-764-877-3776	Sequence 3776, Ap

c 385	30	1.8	562	10	US-09-925-301-599	Sequence 599, App	c 458	30	1.8	32195	10	US-09-764-870-617	Sequence 617, App
c 386	30	1.8	574	9	US-10-040-739-871	Sequence 871, App	c 459	30	1.8	32195	10	US-09-764-869-1605	Sequence 1605, App
c 387	30	1.8	578	10	US-09-764-846-305	Sequence 305, App	c 460	30	1.8	32195	10	US-09-764-869-2017	Sequence 2017, App
c 388	30	1.8	653	10	US-09-764-869-2160	Sequence 2160, App	c 461	30	1.8	32219	10	US-09-764-869-2016	Sequence 2016, App
c 389	30	1.8	661	10	US-09-764-869-3054	Sequence 3054, App	c 462	30	1.8	32249	10	US-09-764-869-2314	Sequence 2314, App
c 390	30	1.8	728	9	US-10-202-193-297	Sequence 297, App	c 463	30	1.8	32249	10	US-09-764-878-202	Sequence 202, App
c 391	30	1.8	742	9	US-10-202-193-296	Sequence 296, App	c 464	30	1.8	35414	10	US-09-836-607-45	Sequence 45, Appl
c 392	30	1.8	768	9	US-10-202-193-136	Sequence 136, App	c 465	30	1.8	38374	10	US-09-880-107-3950	Sequence 3950, App
c 393	30	1.8	772	9	US-10-202-193-152	Sequence 152, App	c 466	30	1.8	38374	10	US-09-880-107-3950	Sequence 3, Appl
c 394	30	1.8	777	9	US-10-202-193-137	Sequence 137, App	c 467	30	1.8	38844	12	US-10-060-333-3	Sequence 3, Appl
c 395	30	1.8	778	9	US-10-202-193-151	Sequence 151, App	c 468	30	1.8	41100	10	US-09-755-665-46	Sequence 46, Appl
c 396	30	1.8	782	9	US-10-202-193-171	Sequence 171, App	c 469	30	1.8	43058	10	US-09-954-456-292	Sequence 292, App
c 397	30	1.8	829	9	US-10-202-193-191	Sequence 191, App	c 470	30	1.8	43058	10	US-09-954-456-529	Sequence 529, App
c 398	30	1.8	1000	10	US-09-784-423-32	Sequence 32, Appl	c 471	30	1.8	43058	10	US-09-880-107-3950	Sequence 3, Appl
c 399	30	1.8	1019	9	US-10-202-193-31	Sequence 31, Appl	c 472	30	1.8	43850	12	US-10-060-332-3	Sequence 5, Appl
c 400	30	1.8	1024	9	US-10-202-193-9	Sequence 9, Appl	c 473	30	1.8	43850	12	US-09-927-091-6	Sequence 6, Appl
c 401	30	1.8	1024	9	US-10-202-193-71	Sequence 71, Appl	c 474	30	1.8	49884	10	US-09-739-457-5	Sequence 2257, App
c 402	30	1.8	1024	9	US-10-202-193-72	Sequence 72, Appl	c 475	30	1.8	49884	10	US-09-954-456-2257	Sequence 18, Appl
c 403	30	1.8	1442	9	US-10-098-841-145	Sequence 145, App	c 476	30	1.8	62944	10	US-09-780-172-18	Sequence 2116, App
c 404	30	1.8	1614	10	US-09-764-860-986	Sequence 986, App	c 477	30	1.8	63000	10	US-09-880-107-3768	Sequence 243, App
c 405	30	1.8	1614	10	US-09-764-860-986	Sequence 987, App	c 478	30	1.8	63000	10	US-09-880-107-3768	Sequence 3, Appl
c 406	30	1.8	1863	9	US-09-974-879-88	Sequence 88, Appl	c 479	30	1.8	65109	10	US-09-854-883-243	Sequence 79, Appl
c 407	30	1.8	1914	10	US-09-745-763-175	Sequence 175, App	c 480	30	1.8	75899	10	US-09-835-232-6	Sequence 7, Appl
c 408	30	1.8	2145	10	US-09-079-892-7	Sequence 7, Appl	c 481	30	1.8	88191	10	US-10-003-806-6	Sequence 6, Appl
c 409	30	1.8	2156	10	US-09-823-901-4	Sequence 4, Appl	c 482	30	1.8	111282	12	US-10-003-806-9	Sequence 9, Appl
c 410	30	1.8	3042	9	US-09-764-869-1801	Sequence 1801, App	c 483	30	1.8	116592	10	US-09-818-512-3	Sequence 3, Appl
c 411	30	1.8	3193	9	US-10-001-857-4	Sequence 4, Appl	c 484	30	1.8	145831	10	US-09-969-708-79	Sequence 3, Appl
c 412	30	1.8	3269	10	US-09-969-347-167	Sequence 167, App	c 485	30	1.8	145831	10	US-09-954-456-2116	Sequence 1, Appl
c 413	30	1.8	3471	10	US-09-764-847-1893	Sequence 1893, App	c 486	30	1.8	170834	10	US-09-764-877-4014	Sequence 4014, App
c 414	30	1.8	3471	10	US-09-764-847-1893	Sequence 1893, App	c 487	30	1.8	170834	10	US-09-764-860-713	Sequence 713, App
c 415	30	1.8	3612	9	US-10-098-841-112	Sequence 112, App	c 488	30	1.8	180216	10	US-09-867-701-1030	Sequence 1030, App
c 416	30	1.8	4660	10	US-09-764-869-1572	Sequence 1572, App	c 489	29	1.7	145	10	US-09-867-701-700	Sequence 700, App
c 417	30	1.8	4857	10	US-09-764-878-268	Sequence 268, App	c 490	29	1.7	153	10	US-09-867-701-9335	Sequence 9335, App
c 418	30	1.8	4857	10	US-09-764-860-947	Sequence 947, App	c 491	29	1.7	235	10	US-09-867-701-126	Sequence 126, App
c 419	30	1.8	5254	10	US-09-764-887-609	Sequence 609, App	c 492	29	1.7	329	9	US-09-878-178-813	Sequence 813, App
c 420	30	1.8	5796	10	US-09-764-878-305	Sequence 305, App	c 493	29	1.7	329	9	US-09-867-701-6521	Sequence 6521, App
c 421	30	1.8	5897	10	US-09-899-569A-1	Sequence 1, Appl	c 494	29	1.7	332	10	US-09-867-701-110	Sequence 110, App
c 422	30	1.8	5914	10	US-09-764-878-303	Sequence 303, App	c 495	29	1.7	332	10	US-09-764-869-568	Sequence 568, App
c 423	30	1.8	6163	10	US-09-899-569A-3	Sequence 3, Appl	c 496	29	1.7	336	10	US-09-867-701-6537	Sequence 6537, App
c 424	30	1.8	6431	10	US-09-954-456-2030	Sequence 2030, App	c 497	29	1.7	345	10	US-09-867-701-2899	Sequence 2899, App
c 425	30	1.8	8858	9	US-09-764-868-1431	Sequence 1431, App	c 498	29	1.7	352	10	US-09-867-701-4511	Sequence 4511, App
c 426	30	1.8	9331	9	US-09-764-847-1647	Sequence 1647, App	c 499	29	1.7	361	10	US-09-998-598-1374	Sequence 9441, App
c 427	30	1.8	10098	9	US-09-764-868-1482	Sequence 1482, App	c 500	29	1.7	366	10	US-09-796-692-4312	Sequence 4312, App
c 428	30	1.8	10138	10	US-09-764-847-1046	Sequence 1046, App	c 501	29	1.7	372	10	US-09-867-701-6435	Sequence 6435, App
c 429	30	1.8	10378	10	US-09-764-847-1938	Sequence 1938, App	c 502	29	1.7	372	10	US-09-796-692-4107	Sequence 4107, App
c 430	30	1.8	12258	10	US-09-801-574-58	Sequence 58, Appl	c 503	29	1.7	373	9	US-09-796-692-9358	Sequence 9358, App
c 431	30	1.8	12452	10	US-09-764-877-3989	Sequence 3989, App	c 504	29	1.7	373	9	US-09-796-692-7035	Sequence 7035, App
c 432	30	1.8	12988	10	US-09-764-869-2285	Sequence 2285, App	c 505	29	1.7	374	9	US-09-796-692-9288	Sequence 9288, App
c 433	30	1.8	13485	10	US-09-764-870-548	Sequence 548, App	c 506	29	1.7	375	9	US-09-867-701-7290	Sequence 7290, App
c 434	30	1.8	13808	10	US-09-764-877-3351	Sequence 3351, App	c 507	29	1.7	382	10	US-09-867-701-7900	Sequence 7900, App
c 435	30	1.8	13873	9	US-09-764-868-1282	Sequence 1282, App	c 508	29	1.7	392	10	US-09-867-701-8517	Sequence 8517, App
c 436	30	1.8	14012	9	US-09-819-994-3	Sequence 3, Appl	c 509	29	1.7	393	10	US-09-764-878-320	Sequence 320, App
c 437	30	1.8	14152	9	US-09-764-869-1573	Sequence 1573, App	c 510	29	1.7	393	10	US-09-946-807-1153	Sequence 321, App
c 438	30	1.8	14175	9	US-09-764-868-1474	Sequence 1474, App	c 511	29	1.7	401	9	US-09-946-807-1153	Sequence 1153, App
c 439	30	1.8	15041	10	US-09-764-869-1421	Sequence 1421, App	c 512	29	1.7	401	9	US-09-946-807-1531	Sequence 1531, App
c 440	30	1.8	15041	10	US-09-764-869-1421	Sequence 1421, App	c 513	29	1.7	401	10	US-09-795-668-1153	Sequence 1153, App
c 441	30	1.8	15745	10	US-09-764-877-3236	Sequence 3236, App	c 514	29	1.7	401	10	US-09-795-668-1471	Sequence 1471, App
c 442	30	1.8	16123	10	US-09-764-877-3053	Sequence 3053, App	c 515	29	1.7	401	10	US-09-795-668-1153	Sequence 1153, App
c 443	30	1.8	18036	10	US-09-764-869-2287	Sequence 2287, App	c 516	29	1.7	403	9	US-09-946-807-1470	Sequence 1470, App
c 444	30	1.8	19369	9	US-10-190-593-3	Sequence 3, Appl	c 517	29	1.7	403	10	US-09-795-668-1472	Sequence 1472, App
c 445	30	1.8	21045	10	US-09-764-864-1695	Sequence 1695, App	c 518	29	1.7	403	10	US-09-795-668-1470	Sequence 1470, App
c 446	30	1.8	21761	10	US-09-764-847-1680	Sequence 1680, App	c 519	29	1.7	403	10	US-09-946-807-1531	Sequence 1531, App
c 447	30	1.8	23580	10	US-09-764-860-990	Sequence 990, App	c 520	29	1.7	403	10	US-09-795-668-1153	Sequence 1153, App
c 448	30	1.8	26928	10	US-09-880-107-2278	Sequence 2278, App	c 521	29	1.7	403	10	US-09-795-668-1471	Sequence 1471, App
c 449	30	1.8	28313	10	US-09-764-877-3194	Sequence 3194, App	c 522	29	1.7	403	10	US-09-795-668-1471	Sequence 1471, App
c 450	30	1.8	28897	10	US-09-764-877-3897	Sequence 3897, App	c 523	29	1.7	403	10	US-09-795-668-1471	Sequence 1471, App
c 451	30	1.8	29228	10	US-09-764-877-3198	Sequence 3198, App	c 524	29	1.7	403	10	US-09-795-668-1471	Sequence 1471, App
c 452	30	1.8	29449	9	US-09-989-442-161	Sequence 161, App	c 525	29	1.7	403	10	US-09-795-668-1471	Sequence 1471, App
c 453	30	1.8	32187	10	US-09-764-847-1550	Sequence 1550, App	c 526	29	1.7	403	10	US-09-795-668-1471	Sequence 1471, App
c 454	30	1.8	32190	10	US-09-764-878-201	Sequence 201, App	c 527	29	1.7	403	10	US-09-795-668-1471	Sequence 1471, App
c 455	30	1.8	32193	10	US-09-764-878-200	Sequence 200, App	c 528	29	1.7	403	10	US-09-795-668-1471	Sequence 1471, App
c 456	30	1.8	32193	10	US-09-764-847-1549	Sequence 1549, App	c 529	29	1.7	403	10	US-09-795-668-1470	Sequence 1470, App
c 457	30	1.8	32195	10	US-09-764-870-611	Sequence 611, App	c 530	29	1.7	403	10	US-09-795-668-1472	Sequence 1472, App

c 531	29	1.7	403	10	US-09-795-686-1470	Sequence 1470, Ap	604	29	1.7	4607	10	US-09-764-878-424	Sequence 424, App
c 532	29	1.7	403	10	US-09-795-686-1472	Sequence 1472, Ap	c 605	29	1.7	5011	10	US-09-764-847-1562	Sequence 1562, Ap
c 533	29	1.7	417	10	US-09-867-701-8286	Sequence 8286, Ap	c 606	29	1.7	5088	10	US-09-764-869-1760	Sequence 1760, Ap
c 534	29	1.7	425	9	US-09-292-758-7	Sequence 7, Appl1	c 607	29	1.7	5150	9	US-09-764-904-94	Sequence 94, Appl1
c 535	29	1.7	425	10	US-09-867-701-8101	Sequence 8101, Ap	c 608	29	1.7	5150	10	US-09-764-860-1138	Sequence 1138, Ap
c 536	29	1.7	437	10	US-09-867-701-8896	Sequence 8896, Ap	c 609	29	1.7	5217	10	US-09-764-877-2963	Sequence 2963, Ap
c 537	29	1.7	437	10	US-09-867-701-9057	Sequence 9057, Ap	c 610	29	1.7	5217	10	US-09-764-877-2964	Sequence 2964, Ap
c 538	29	1.7	442	10	US-09-867-701-153	Sequence 153, App	c 611	29	1.7	5238	10	US-09-764-877-2096	Sequence 2096, Ap
c 539	29	1.7	444	10	US-09-867-701-6812	Sequence 6812, Ap	c 612	29	1.7	5770	10	US-09-908-711-144	Sequence 144, App
c 540	29	1.7	455	9	US-09-796-692-8498	Sequence 8498, Ap	c 613	29	1.7	6074	10	US-09-880-107-2378	Sequence 2378, Ap
c 541	29	1.7	459	10	US-09-867-701-5927	Sequence 5927, Ap	c 614	29	1.7	6149	10	US-09-764-887-436	Sequence 436, App
c 542	29	1.7	461	10	US-09-867-701-2179	Sequence 2179, Ap	c 615	29	1.7	6149	10	US-09-764-887-437	Sequence 437, App
c 543	29	1.7	471	9	US-09-822-846-457	Sequence 457, App	c 616	29	1.7	6250	10	US-09-764-877-2965	Sequence 2965, Ap
c 544	29	1.7	473	10	US-09-880-107-361	Sequence 361, App	c 617	29	1.7	6515	10	US-09-764-877-2635	Sequence 2635, Ap
c 545	29	1.7	494	10	US-09-867-701-6775	Sequence 6775, Ap	c 618	29	1.7	6533	12	US-10-044-090-531	Sequence 531, App
c 546	29	1.7	506	10	US-09-867-701-10583	Sequence 10583, A	c 619	29	1.7	6682	10	US-09-764-864-1699	Sequence 1699, Ap
c 547	29	1.7	587	10	US-09-864-761-9867	Sequence 9867, Ap	c 620	29	1.7	6912	10	US-09-764-869-2354	Sequence 2354, Ap
c 548	29	1.7	590	10	US-09-764-847-1070	Sequence 1070, Ap	c 621	29	1.7	7017	10	US-09-764-877-3773	Sequence 3773, Ap
c 549	29	1.7	601	10	US-09-777-921A-82	Sequence 82, Appl	c 622	29	1.7	7032	10	US-09-764-847-1429	Sequence 1429, Ap
c 550	29	1.7	601	10	US-09-777-921A-83	Sequence 83, Appl	c 623	29	1.7	7373	10	US-09-764-853-896	Sequence 896, App
c 551	29	1.7	761	9	US-09-764-904-108	Sequence 108, App	c 624	29	1.7	7606	10	US-09-764-860-800	Sequence 800, App
c 552	29	1.7	761	10	US-09-764-860-1163	Sequence 1163, Ap	c 625	29	1.7	7608	10	US-09-764-860-801	Sequence 801, App
c 553	29	1.7	762	12	US-10-001-843-72	Sequence 72, Appl	c 626	29	1.7	7683	10	US-09-764-847-1173	Sequence 1173, App
c 554	29	1.7	823	9	US-10-202-193-216	Sequence 216, App	c 627	29	1.7	7739	10	US-09-764-877-3189	Sequence 3189, Ap
c 555	29	1.7	840	9	US-09-764-868-65	Sequence 65, Appl	c 628	29	1.7	7739	10	US-09-764-864-1723	Sequence 1723, Ap
c 556	29	1.7	919	10	US-09-764-887-518	Sequence 518, App	c 629	29	1.7	7960	10	US-09-764-869-2327	Sequence 2327, Ap
c 557	29	1.7	944	10	US-09-834-975-880	Sequence 880, App	c 630	29	1.7	8095	10	US-09-880-107-2228	Sequence 2228, Ap
c 558	29	1.7	946	10	US-09-834-975-881	Sequence 881, App	c 631	29	1.7	8121	10	US-09-764-869-1689	Sequence 1689, Ap
c 559	29	1.7	946	9	US-09-974-879-136	Sequence 136, App	c 632	29	1.7	8269	10	US-09-764-887-452	Sequence 452, App
c 560	29	1.7	958	9	US-09-974-879-81	Sequence 81, Appl	c 633	29	1.7	8269	10	US-09-764-870-551	Sequence 551, App
c 561	29	1.7	966	10	US-09-822-849A-218	Sequence 218, App	c 634	29	1.7	8269	10	US-09-764-853-879	Sequence 879, App
c 562	29	1.7	989	12	US-10-001-843-73	Sequence 73, Appl	c 635	29	1.7	8746	10	US-09-764-860-1022	Sequence 1022, Ap
c 563	29	1.7	998	9	US-09-983-802-62	Sequence 62, Appl	c 636	29	1.7	8758	10	US-09-962-276-3	Sequence 3, Appl1
c 564	29	1.7	1014	9	US-09-729-835-32	Sequence 32, Appl	c 637	29	1.7	8821	9	US-09-764-868-1497	Sequence 1497, App
c 565	29	1.7	1031	9	US-10-152-661-470	Sequence 470, Appl	c 638	29	1.7	8891	10	US-09-764-854-1657	Sequence 1657, Ap
c 566	29	1.7	1053	10	US-09-729-835-31	Sequence 31, Appl	c 639	29	1.7	8895	10	US-09-764-855-332	Sequence 332, App
c 567	29	1.7	1076	9	US-10-152-661-467	Sequence 467, Appl	c 640	29	1.7	8896	10	US-09-764-855-331	Sequence 331, App
c 568	29	1.7	1137	10	US-09-995-494-21	Sequence 21, Appl	c 641	29	1.7	8932	10	US-09-764-877-3049	Sequence 3049, Ap
c 569	29	1.7	1200	10	US-09-784-423-37	Sequence 37, Appl	c 642	29	1.7	9372	10	US-09-764-877-3244	Sequence 3244, Ap
c 570	29	1.7	1473	10	US-09-822-849A-332	Sequence 332, App	c 643	29	1.7	9439	10	US-09-764-877-2324	Sequence 2324, Ap
c 571	29	1.7	1579	10	US-09-764-878-423	Sequence 423, App	c 644	29	1.7	9474	10	US-09-764-878-372	Sequence 372, App
c 572	29	1.7	1579	10	US-09-764-878-425	Sequence 425, App	c 645	29	1.7	9513	10	US-09-764-869-1811	Sequence 1811, Ap
c 573	29	1.7	1656	10	US-09-955-494-17	Sequence 17, Appl	c 646	29	1.7	9669	10	US-09-764-869-2072	Sequence 2072, Ap
c 574	29	1.7	1664	10	US-09-764-877-2841	Sequence 2841, Ap	c 647	29	1.7	9887	9	US-10-000-639-7	Sequence 7, Appl1
c 575	29	1.7	1936	10	US-09-764-877-2378	Sequence 2378, Ap	c 648	29	1.7	9903	10	US-09-764-855-262	Sequence 262, App
c 576	29	1.7	2027	10	US-09-731-872-77	Sequence 77, Appl	c 649	29	1.7	9970	10	US-09-764-877-2277	Sequence 2277, Ap
c 577	29	1.7	2109	10	US-09-764-855-325	Sequence 325, App	c 650	29	1.7	10503	10	US-09-764-870-579	Sequence 579, App
c 578	29	1.7	2147	10	US-09-764-847-1738	Sequence 1738, Ap	c 651	29	1.7	10503	10	US-09-764-853-854	Sequence 854, App
c 579	29	1.7	2260	9	US-09-822-846-247	Sequence 247, App	c 652	29	1.7	10503	10	US-09-764-860-1032	Sequence 1032, Ap
c 580	29	1.7	2351	9	US-09-798-889-37	Sequence 37, Appl	c 653	29	1.7	11057	9	US-09-764-860-1308	Sequence 1308, Ap
c 581	29	1.7	2380	10	US-09-764-877-3177	Sequence 3177, Ap	c 654	29	1.7	11427	10	US-09-764-869-1812	Sequence 1812, Ap
c 582	29	1.7	2394	10	US-09-764-877-3176	Sequence 3176, Ap	c 655	29	1.7	12127	10	US-09-764-860-1113	Sequence 1113, Ap
c 583	29	1.7	2403	10	US-09-880-107-3343	Sequence 3343, Ap	c 656	29	1.7	12149	10	US-09-764-869-2258	Sequence 2258, Ap
c 584	29	1.7	2532	10	US-09-962-276-1	Sequence 1, Appl1	c 657	29	1.7	12259	10	US-09-764-869-1690	Sequence 1690, Ap
c 585	29	1.7	2533	10	US-09-764-869-1656	Sequence 1656, Ap	c 658	29	1.7	12493	10	US-09-764-864-1715	Sequence 1715, Ap
c 586	29	1.7	2687	10	US-09-764-869-2099	Sequence 2099, Ap	c 659	29	1.7	12542	10	US-09-764-864-1774	Sequence 1774, Ap
c 587	29	1.7	2687	10	US-09-764-869-2100	Sequence 2100, Ap	c 660	29	1.7	12566	10	US-09-764-869-2035	Sequence 2035, Ap
c 588	29	1.7	2923	10	US-09-837-446-5	Sequence 5, Appl1	c 661	29	1.7	12932	10	US-09-764-847-1132	Sequence 1132, Ap
c 589	29	1.7	2972	10	US-09-779-239-1	Sequence 1, Appl1	c 662	29	1.7	13058	10	US-09-764-846-303	Sequence 303, App
c 590	29	1.7	3144	9	US-10-163-866-31	Sequence 31, Appl	c 663	29	1.7	13176	10	US-09-764-869-1657	Sequence 1657, Ap
c 591	29	1.7	3459	10	US-09-764-887-412	Sequence 412, App	c 664	29	1.7	13224	10	US-09-764-853-897	Sequence 897, App
c 592	29	1.7	3460	10	US-09-739-254-44	Sequence 44, Appl	c 665	29	1.7	13273	10	US-09-764-869-2349	Sequence 2349, Ap
c 593	29	1.7	3460	10	US-09-904-615-44	Sequence 44, Appl	c 666	29	1.7	13273	10	US-09-764-869-2349	Sequence 2349, Ap
c 594	29	1.7	3512	10	US-09-764-847-1974	Sequence 1974, Ap	c 667	29	1.7	13467	9	US-09-764-868-1330	Sequence 1330, Ap
c 595	29	1.7	3560	10	US-09-764-869-2411	Sequence 2411, Ap	c 668	29	1.7	13467	9	US-09-764-868-1330	Sequence 1330, Ap
c 596	29	1.7	3561	10	US-09-764-869-2410	Sequence 2410, Ap	c 669	29	1.7	13485	10	US-09-764-870-548	Sequence 548, App
c 597	29	1.7	3694	10	US-09-880-107-1616	Sequence 1616, Ap	c 670	29	1.7	13608	9	US-09-764-868-1369	Sequence 1369, Ap
c 598	29	1.7	3699	10	US-09-822-849A-14	Sequence 14, Appl	c 671	29	1.7	13608	10	US-09-764-877-3351	Sequence 3351, Ap
c 599	29	1.7	3854	10	US-09-764-877-3882	Sequence 3882, Ap	c 672	29	1.7	14012	9	US-09-819-994-3	Sequence 3, Appl1
c 600	29	1.7	4150	12	US-10-016-725-15	Sequence 15, Appl	c 673	29	1.7	14598	9	US-09-764-868-1429	Sequence 1429, Ap
c 601	29	1.7	4346	12	US-10-044-090-296	Sequence 296, App	c 674	29	1.7	15016	10	US-09-880-107-3783	Sequence 3783, Ap
c 602	29	1.7	4485	10	US-09-764-877-5981	Sequence 5981, Ap	c 675	29	1.7	15016	10	US-09-880-107-3783	Sequence 3783, Ap
c 603	29	1.7	4607	10	US-09-764-878-422	Sequence 422, App	c 676	29	1.7	15070	10	US-09-764-877-3979	Sequence 3979, Ap

677	29	1.7	15071	10	US-09-764-877-3978	Sequence 3378, Ap	750	29	1.7	32193	10	US-09-764-877-2623	Sequence 2623, Ap
c 678	29	1.7	15287	10	US-09-764-877-2134	Sequence 2134, Ap	c 751	29	1.7	32193	10	US-09-764-877-2623	Sequence 2623, Ap
679	29	1.7	15714	10	US-09-764-869-2355	Sequence 2355, Ap	c 752	29	1.7	32195	10	US-09-764-847-1512	Sequence 1512, Ap
680	29	1.7	16181	10	US-09-764-847-1426	Sequence 1426, Ap	c 753	29	1.7	32204	10	US-09-764-846-306	Sequence 306, App
681	29	1.7	16337	10	US-09-764-877-3468	Sequence 3468, Ap	c 754	29	1.7	32204	10	US-09-764-855-327	Sequence 327, App
682	29	1.7	16552	10	US-09-764-855-321	Sequence 321, App	c 755	29	1.7	32221	10	US-09-764-847-1406	Sequence 1406, Ap
683	29	1.7	16552	10	US-09-764-855-322	Sequence 322, App	c 756	29	1.7	32221	10	US-09-764-847-1406	Sequence 1406, Ap
c 684	29	1.7	16854	10	US-09-764-878-217	Sequence 217, App	c 757	29	1.7	32248	10	US-09-764-860-802	Sequence 802, App
c 685	29	1.7	17239	10	US-09-764-869-2326	Sequence 2326, Ap	c 758	29	1.7	32249	10	US-09-764-878-202	Sequence 202, App
c 686	29	1.7	17379	10	US-09-764-860-861	Sequence 861, App	c 759	29	1.7	32023	10	US-09-880-107-3350	Sequence 3350, Ap
c 687	29	1.7	17498	10	US-09-764-860-798	Sequence 798, App	c 760	29	1.7	36159	12	US-10-135-687-3	Sequence 3, Appli
c 688	29	1.7	17509	10	US-09-880-107-2097	Sequence 2097, Ap	c 761	29	1.7	36651	10	US-09-964-469-3	Sequence 3, Appli
c 689	29	1.7	17509	10	US-09-748-127-3	Sequence 3, Appli	c 762	29	1.7	36741	9	US-09-782-378A-12	Sequence 12, Appli
c 690	29	1.7	17792	10	US-09-764-869-1599	Sequence 1599, Ap	c 763	29	1.7	36776	9	US-10-160-293-3	Sequence 3, Appli
c 691	29	1.7	18400	10	US-09-901-151-3	Sequence 3, Appli	c 764	29	1.7	41907	10	US-09-967-0113-5	Sequence 5, Appli
c 692	29	1.7	19553	10	US-09-764-847-1425	Sequence 1425, Ap	c 765	29	1.7	44848	9	US-09-988-113-42	Sequence 42, Appli
c 693	29	1.7	20210	10	US-09-764-870-598	Sequence 598, App	c 766	29	1.7	48848	10	US-09-776-874A-42	Sequence 42, Appli
c 694	29	1.7	20272	10	US-09-908-711-145	Sequence 145, App	c 767	29	1.7	48939	12	US-10-025-187-3	Sequence 3, Appli
c 695	29	1.7	20522	10	US-09-764-877-3774	Sequence 3774, Ap	c 768	29	1.7	46718	10	US-09-816-093-3	Sequence 3, Appli
696	29	1.7	20966	10	US-09-776-976-7	Sequence 7, Appli	c 769	29	1.7	49984	10	US-09-739-457-5	Sequence 5, Appli
697	29	1.7	20966	10	US-09-758-055-7	Sequence 7, Appli	c 770	29	1.7	52216	10	US-09-747-810-1	Sequence 1, Appli
c 698	29	1.7	20966	10	US-09-909-547-7	Sequence 7, Appli	c 771	29	1.7	53226	10	US-09-818-264-3	Sequence 3, Appli
c 699	29	1.7	21129	10	US-09-764-869-1734	Sequence 1734, Ap	c 772	29	1.7	53542	10	US-09-801-574-61	Sequence 61, Appli
c 700	29	1.7	21784	10	US-09-820-002-3	Sequence 3, Appli	c 773	29	1.7	62804	12	US-10-096-960-3	Sequence 3, Appli
c 701	29	1.7	21860	10	US-09-784-877-3803	Sequence 3803, Ap	c 774	29	1.7	65042	9	US-10-229-124-3	Sequence 3, Appli
702	29	1.7	22255	9	US-09-976-740-51	Sequence 51, Appli	c 775	29	1.7	65608	9	US-09-954-531-180	Sequence 180, App
703	29	1.7	22255	12	US-10-023-529-51	Sequence 51, Appli	c 776	29	1.7	65608	10	US-09-962-436-292	Sequence 292, App
704	29	1.7	22255	12	US-10-023-523-51	Sequence 51, Appli	c 777	29	1.7	65608	10	US-09-962-832-119	Sequence 119, App
c 705	29	1.7	22786	10	US-09-764-877-3727	Sequence 3727, Ap	c 778	29	1.7	66686	10	US-09-736-960-86	Sequence 86, Appli
c 706	29	1.7	22889	9	US-09-860-670-164	Sequence 164, App	c 779	29	1.7	66804	10	US-09-740-041-3	Sequence 3, Appli
707	29	1.7	23433	10	US-09-927-091-7	Sequence 7, Appli	c 780	29	1.7	68004	10	US-09-740-041-3	Sequence 3, Appli
708	29	1.7	23934	10	US-09-764-860-777	Sequence 777, App	c 781	29	1.7	80959	9	US-09-858-546-3	Sequence 3, Appli
c 709	29	1.7	23934	10	US-09-764-877-2536	Sequence 2536, Ap	c 782	29	1.7	98865	10	US-09-770-689A-3	Sequence 3, Appli
710	29	1.7	23934	10	US-09-764-877-2544	Sequence 2544, Ap	c 783	29	1.7	98865	10	US-09-770-689A-3	Sequence 3, Appli
711	29	1.7	24533	9	US-09-764-868-1349	Sequence 1349, Ap	c 784	29	1.7	99014	10	US-09-880-107-3428	Sequence 3428, Ap
712	29	1.7	24768	10	US-09-764-887-602	Sequence 602, App	c 785	29	1.7	110096	10	US-09-880-107-1542	Sequence 1542, Ap
713	29	1.7	24768	10	US-09-764-855-193	Sequence 193, App	c 786	29	1.7	110096	10	US-09-880-107-1542	Sequence 1542, Ap
714	29	1.7	25424	10	US-09-764-855-194	Sequence 194, App	c 787	29	1.7	116592	10	US-09-818-512-3	Sequence 3, Appli
715	29	1.7	25950	10	US-09-764-870-597	Sequence 597, App	c 788	29	1.7	143306	10	US-09-729-920-3	Sequence 3, Appli
716	29	1.7	26006	10	US-09-764-869-1963	Sequence 1963, Ap	c 789	29	1.7	148567	9	US-10-254-869-3	Sequence 3, Appli
717	29	1.7	26006	10	US-09-764-864-1638	Sequence 1638, Ap	c 790	29	1.7	148567	10	US-09-801-876B-3	Sequence 3, Appli
c 718	29	1.7	26013	10	US-09-764-869-1961	Sequence 1961, Ap	c 791	29	1.7	152331	9	US-10-095-407-16	Sequence 16, Appli
c 719	29	1.7	26013	10	US-09-764-864-1636	Sequence 1636, Ap	c 792	29	1.7	167343	10	US-09-962-436-281	Sequence 281, App
720	29	1.7	26018	10	US-09-764-869-1962	Sequence 1962, Ap	c 793	29	1.7	167343	10	US-09-964-824A-273	Sequence 273, App
721	29	1.7	26018	10	US-09-764-864-1637	Sequence 1637, Ap	c 794	29	1.7	172637	10	US-09-805-458A-3	Sequence 3, Appli
722	29	1.7	26657	10	US-09-810-673A-3	Sequence 3, Appli	c 795	29	1.7	174493	9	US-10-238-709-3	Sequence 3, Appli
723	29	1.7	26657	10	US-09-962-832-222	Sequence 222, App	c 796	29	1.7	174493	10	US-09-804-471A-3	Sequence 3, Appli
724	29	1.7	26928	10	US-09-880-107-2278	Sequence 2278, Ap	c 797	29	1.7	176373	9	US-10-095-407-17	Sequence 17, Appli
c 725	29	1.7	26995	10	US-09-764-878-227	Sequence 227, App	c 798	29	1.7	180216	10	US-09-835-232-6	Sequence 6, Appli
c 726	29	1.7	27377	10	US-09-816-248-18	Sequence 18, Appli	c 799	29	1.7	197496	9	US-09-877-177-10	Sequence 10, Appli
727	29	1.7	29695	10	US-09-752-820A-3	Sequence 3, Appli	c 800	28	1.7	302250	10	US-09-962-832-154	Sequence 154, App
728	29	1.7	29695	10	US-09-813-319A-3	Sequence 3, Appli	c 801	28	1.7	138	9	US-09-764-868-1359	Sequence 1359, Ap
c 729	29	1.7	29921	9	US-10-083-853-1	Sequence 1, Appli	c 802	28	1.7	158	10	US-09-764-860-1049	Sequence 1049, Ap
730	29	1.7	30625	10	US-09-927-091-5	Sequence 5, Appli	c 803	28	1.7	158	10	US-09-867-701-7069	Sequence 7069, Ap
731	29	1.7	30676	10	US-09-927-091-8	Sequence 8, Appli	c 804	28	1.7	198	10	US-09-867-701-7069	Sequence 7069, Ap
c 732	29	1.7	31208	10	US-09-852-067-3	Sequence 3, Appli	c 805	28	1.7	208	10	US-09-764-847-1017	Sequence 1017, Ap
c 733	29	1.7	31314	10	US-09-817-877-3875	Sequence 3875, Ap	c 806	28	1.7	222	10	US-09-867-701-8843	Sequence 8843, Ap
c 734	29	1.7	31814	10	US-09-817-182-3	Sequence 3, Appli	c 807	28	1.7	227	10	US-09-867-701-8190	Sequence 8190, Ap
735	29	1.7	31871	10	US-09-764-847-1403	Sequence 3, Appli	c 808	28	1.7	238	10	US-09-764-887-539	Sequence 539, App
c 736	29	1.7	32146	10	US-09-764-860-797	Sequence 797, App	c 808	28	1.7	238	10	US-09-764-869-2107	Sequence 2107, Ap
c 737	29	1.7	32152	10	US-09-764-855-328	Sequence 328, App	c 809	28	1.7	239	10	US-09-867-701-8426	Sequence 8426, Ap
c 738	29	1.7	32169	10	US-09-764-847-1963	Sequence 232, App	c 810	28	1.7	240	10	US-09-783-590-5759	Sequence 5759, Ap
c 739	29	1.7	32174	9	US-09-860-670-232	Sequence 232, App	c 811	28	1.7	245	9	US-10-040-739-1286	Sequence 1286, Ap
c 740	29	1.7	32174	9	US-09-764-904-90	Sequence 90, Appli	c 812	28	1.7	246	10	US-09-764-855-297	Sequence 297, App
c 741	29	1.7	32174	10	US-09-908-711-158	Sequence 158, App	c 813	28	1.7	270	10	US-09-783-590-2927	Sequence 2927, Ap
c 742	29	1.7	32174	10	US-09-764-860-1134	Sequence 1134, Ap	c 814	28	1.7	298	9	US-10-040-739-1288	Sequence 1288, Ap
743	29	1.7	32174	10	US-09-764-877-2645	Sequence 2645, Ap	c 815	28	1.7	303	10	US-09-867-701-8629	Sequence 8629, Ap
744	29	1.7	32183	10	US-09-764-869-1494	Sequence 1494, Ap	c 816	28	1.7	332	9	US-09-796-692-6817	Sequence 6817, Ap
745	29	1.7	32190	10	US-09-764-887-338	Sequence 338, App	c 817	28	1.7	332	10	US-09-867-701-8076	Sequence 8076, Ap
c 746	29	1.7	32190	10	US-09-764-878-201	Sequence 201, App	c 818	28	1.7	344	10	US-09-867-701-6382	Sequence 6382, Ap
747	29	1.7	32192	10	US-09-764-877-2844	Sequence 2844, Ap	c 819	28	1.7	342	10	US-09-867-701-7760	Sequence 7760, Ap
c 748	29	1.7	32192	10	US-09-764-877-2844	Sequence 2844, Ap	c 820	28	1.7	354	10	US-09-867-701-8056	Sequence 8056, Ap
c 749	29	1.7	32193	10	US-09-764-878-200	Sequence 200, App	c 821	28	1.7	365	10	US-09-867-701-7234	Sequence 7234, Ap
	29	1.7	32193	10	US-09-764-878-200	Sequence 200, App	c 822	28	1.7	365	10	US-09-867-701-7234	Sequence 7234, Ap

c 823	28	1.7	367	10	US-09-867-701-8012	Sequence 8012, Ap	c 896	28	1.7	586	10	US-09-764-869-2281	Sequence 2281, Ap
c 824	28	1.7	370	10	US-09-880-107-1880	Sequence 1880, Ap	c 897	28	1.7	593	10	US-09-764-869-2278	Sequence 2278, Ap
c 825	28	1.7	374	10	US-09-969-347-1113	Sequence 1113, App	c 898	28	1.7	593	10	US-09-764-869-2280	Sequence 2280, Ap
c 826	28	1.7	374	10	US-09-880-107-1402	Sequence 1402, Ap	c 899	28	1.7	600	10	US-09-864-761-7790	Sequence 7790, Ap
c 827	28	1.7	381	10	US-09-815-343-206	Sequence 206, App	c 900	28	1.7	601	10	US-09-818-656A-55	Sequence 55, Appl
c 828	28	1.7	386	9	US-09-954-531-1316	Sequence 1316, Ap	c 901	28	1.7	601	10	US-09-818-656A-56	Sequence 56, Appl
c 829	28	1.7	387	10	US-09-960-352-7819	Sequence 7819, Ap	c 902	28	1.7	652	10	US-09-879-536-717	Sequence 717, App
c 830	28	1.7	391	10	US-09-867-701-7901	Sequence 7901, Ap	c 903	28	1.7	661	9	US-09-822-846-529	Sequence 529, App
c 831	28	1.7	392	10	US-09-867-701-9069	Sequence 9069, Ap	c 904	28	1.7	697	10	US-09-833-381-10	Sequence 10, Appl
c 832	28	1.7	394	10	US-09-867-701-7325	Sequence 7325, Ap	c 905	28	1.7	716	9	US-10-001-835-88	Sequence 88, Appl
c 833	28	1.7	394	10	US-09-867-701-7336	Sequence 7336, Ap	c 906	28	1.7	719	9	US-09-983-802-74	Sequence 74, Appl
c 834	28	1.7	395	10	US-09-867-701-5803	Sequence 5803, Ap	c 907	28	1.7	729	9	US-10-001-857-81	Sequence 81, Appl
c 835	28	1.7	399	10	US-09-880-107-2480	Sequence 2480, Ap	c 908	28	1.7	792	10	US-09-764-898-13	Sequence 13, Appl
c 836	28	1.7	401	9	US-09-946-807-324	Sequence 324, App	c 909	28	1.7	808	9	US-10-040-739-865	Sequence 865, Appl
c 837	28	1.7	401	9	US-09-946-807-832	Sequence 832, App	c 910	28	1.7	846	9	US-10-001-835-96	Sequence 96, Appl
c 838	28	1.7	401	9	US-09-946-807-937	Sequence 937, App	c 911	28	1.7	858	10	US-09-764-877-850	Sequence 850, App
c 839	28	1.7	401	9	US-09-946-807-1087	Sequence 1087, Ap	c 912	28	1.7	879	9	US-09-924-340-73	Sequence 73, Appl
c 840	28	1.7	401	10	US-09-795-668-324	Sequence 324, App	c 913	28	1.7	879	9	US-09-992-600A-73	Sequence 73, Appl
c 841	28	1.7	401	10	US-09-795-668-832	Sequence 832, App	c 914	28	1.7	884	10	US-09-764-853-64	Sequence 64, Appl
c 842	28	1.7	401	10	US-09-795-668-937	Sequence 937, App	c 915	28	1.7	957	9	US-09-986-480-56	Sequence 56, Appl
c 843	28	1.7	401	10	US-09-795-668-1087	Sequence 1087, Ap	c 916	28	1.7	997	10	US-09-764-877-2210	Sequence 2210, Ap
c 844	28	1.7	401	10	US-09-795-668-334	Sequence 324, App	c 917	28	1.7	1101	10	US-09-731-873-54	Sequence 54, Appl
c 845	28	1.7	401	10	US-09-795-668-832	Sequence 832, App	c 918	28	1.7	1104	10	US-09-764-853-266	Sequence 266, App
c 846	28	1.7	401	10	US-09-795-668-937	Sequence 937, App	c 919	28	1.7	1211	10	US-09-891-171-2	Sequence 2, Appl
c 847	28	1.7	401	10	US-09-795-668-1087	Sequence 1087, Ap	c 920	28	1.7	1204	10	US-09-822-849A-22	Sequence 22, Appl
c 848	28	1.7	402	10	US-09-867-701-7868	Sequence 7868, Ap	c 921	28	1.7	1251	10	US-09-822-830A-317	Sequence 317, App
c 849	28	1.7	407	10	US-09-815-343-441	Sequence 441, App	c 922	28	1.7	1257	10	US-09-764-869-541	Sequence 541, App
c 850	28	1.7	407	10	US-09-867-701-8568	Sequence 8568, Ap	c 923	28	1.7	1264	9	US-09-822-846-154	Sequence 154, App
c 851	28	1.7	414	9	US-09-946-807-1403	Sequence 1403, Ap	c 924	28	1.7	1284	10	US-09-764-864-1742	Sequence 1742, Ap
c 852	28	1.7	414	10	US-09-795-668-1403	Sequence 1403, Ap	c 925	28	1.7	1334	9	US-10-174-590-339	Sequence 339, App
c 853	28	1.7	414	10	US-09-795-668-1403	Sequence 1403, Ap	c 926	28	1.7	1334	9	US-10-176-758-339	Sequence 339, App
c 854	28	1.7	415	10	US-09-793-590-11905	Sequence 11905, A	c 927	28	1.7	1334	9	US-10-175-737-339	Sequence 339, App
c 855	28	1.7	417	10	US-09-867-701-10648	Sequence 10648, A	c 928	28	1.7	1334	9	US-10-175-706-339	Sequence 339, App
c 856	28	1.7	424	10	US-09-867-701-10648	Sequence 10648, A	c 929	28	1.7	1334	9	US-10-175-738-339	Sequence 339, App
c 857	28	1.7	424	9	US-09-867-701-5629	Sequence 5629, Ap	c 930	28	1.7	1334	9	US-10-175-752-339	Sequence 339, App
c 858	28	1.7	430	9	US-09-764-868-1302	Sequence 1025, Ap	c 931	28	1.7	1334	9	US-10-176-482-339	Sequence 339, App
c 859	28	1.7	430	9	US-09-764-868-1302	Sequence 1302, Ap	c 932	28	1.7	1334	9	US-10-176-757-339	Sequence 339, App
c 860	28	1.7	430	10	US-09-867-701-10152	Sequence 10152, A	c 933	28	1.7	1334	9	US-10-176-913-339	Sequence 339, App
c 861	28	1.7	434	10	US-09-867-701-8920	Sequence 8920, Ap	c 934	28	1.7	1334	9	US-10-180-552-339	Sequence 339, App
c 862	28	1.7	436	10	US-09-867-701-4136	Sequence 4136, Ap	c 935	28	1.7	1334	9	US-10-180-557-339	Sequence 339, App
c 863	28	1.7	436	10	US-09-867-701-8582	Sequence 8582, Ap	c 936	28	1.7	1334	9	US-10-173-700-339	Sequence 339, App
c 864	28	1.7	440	10	US-09-867-701-7790	Sequence 7790, Ap	c 937	28	1.7	1334	9	US-10-174-572-339	Sequence 339, App
c 865	28	1.7	441	9	US-09-796-692-7801	Sequence 7801, Ap	c 938	28	1.7	1334	9	US-10-174-579-339	Sequence 339, App
c 866	28	1.7	441	10	US-09-867-701-277	Sequence 277, App	c 939	28	1.7	1334	9	US-10-174-582-339	Sequence 339, App
c 867	28	1.7	453	10	US-09-867-701-6806	Sequence 6806, Ap	c 940	28	1.7	1334	9	US-10-174-588-339	Sequence 339, App
c 868	28	1.7	458	10	US-09-867-701-2993	Sequence 2993, Ap	c 941	28	1.7	1334	9	US-10-175-739-339	Sequence 339, App
c 869	28	1.7	463	10	US-09-867-701-10007	Sequence 10007, A	c 942	28	1.7	1334	9	US-10-175-740-339	Sequence 339, App
c 870	28	1.7	464	10	US-09-764-878-402	Sequence 402, App	c 943	28	1.7	1334	9	US-10-175-743-339	Sequence 339, App
c 871	28	1.7	468	10	US-09-867-701-9386	Sequence 9386, Ap	c 944	28	1.7	1334	9	US-10-176-488-339	Sequence 339, App
c 872	28	1.7	470	10	US-09-880-107-454	Sequence 454, App	c 945	28	1.7	1334	9	US-10-176-492-339	Sequence 339, App
c 873	28	1.7	476	10	US-09-867-701-3633	Sequence 3633, Ap	c 946	28	1.7	1334	9	US-10-176-747-339	Sequence 339, App
c 874	28	1.7	486	10	US-09-867-701-7152	Sequence 7152, Ap	c 947	28	1.7	1334	9	US-10-176-750-339	Sequence 339, App
c 875	28	1.7	488	10	US-09-867-701-6777	Sequence 6777, Ap	c 948	28	1.7	1334	9	US-10-176-985-339	Sequence 339, App
c 876	28	1.7	489	9	US-09-854-133-109	Sequence 109, App	c 949	28	1.7	1334	9	US-10-176-987-339	Sequence 339, App
c 877	28	1.7	489	10	US-09-738-973-109	Sequence 109, App	c 950	28	1.7	1334	9	US-10-176-991-339	Sequence 339, App
c 878	28	1.7	491	10	US-09-867-701-6700	Sequence 6700, Ap	c 951	28	1.7	1334	9	US-10-176-992-339	Sequence 339, App
c 879	28	1.7	492	9	US-09-796-692-3679	Sequence 3679, Ap	c 952	28	1.7	1334	9	US-10-176-993-339	Sequence 339, App
c 880	28	1.7	498	10	US-09-920-300A-1487	Sequence 1487, Ap	c 953	28	1.7	1334	9	US-10-176-993-339	Sequence 339, App
c 881	28	1.7	498	12	US-10-033-528-1487	Sequence 1487, Ap	c 954	28	1.7	1334	12	US-10-184-658-339	Sequence 339, App
c 882	28	1.7	500	10	US-09-827-998-59	Sequence 59, Appl	c 955	28	1.7	1334	10	US-10-052-586-339	Sequence 339, App
c 883	28	1.7	501	10	US-09-867-701-1298	Sequence 1298, Ap	c 956	28	1.7	1351	10	US-09-731-872-80	Sequence 80, Appl
c 884	28	1.7	506	10	US-09-817-190A-14	Sequence 14, Appl	c 957	28	1.7	1359	10	US-09-925-299-200	Sequence 200, App
c 885	28	1.7	510	10	US-09-783-590-5651	Sequence 5651, Ap	c 958	28	1.7	1399	9	US-09-989-293A-27	Sequence 27, Appl
c 886	28	1.7	517	10	US-09-867-701-876	Sequence 876, Ap	c 959	28	1.7	1399	9	US-09-989-735-27	Sequence 27, Appl
c 887	28	1.7	534	9	US-10-046-935-1395	Sequence 1395, Ap	c 960	28	1.7	1399	9	US-09-990-444-27	Sequence 27, Appl
c 888	28	1.7	534	9	US-09-878-178-1395	Sequence 1395, Ap	c 961	28	1.7	1399	9	US-09-989-730-27	Sequence 27, Appl
c 889	28	1.7	540	9	US-09-796-692-7184	Sequence 7184, Ap	c 962	28	1.7	1399	9	US-09-990-436-27	Sequence 27, Appl
c 890	28	1.7	567	10	US-09-764-847-1069	Sequence 1069, Ap	c 963	28	1.7	1399	9	US-09-991-181-27	Sequence 27, Appl
c 891	28	1.7	572	9	US-09-796-692-8715	Sequence 8715, Ap	c 964	28	1.7	1399	9	US-09-993-687-27	Sequence 27, Appl
c 892	28	1.7	575	10	US-09-764-847-1841	Sequence 1841, Ap	c 965	28	1.7	1399	9	US-09-989-734-27	Sequence 27, Appl
c 893	28	1.7	576	10	US-09-764-847-1071	Sequence 1071, Ap	c 966	28	1.7	1399	9	US-09-997-653-27	Sequence 27, Appl
c 894	28	1.7	576	10	US-09-764-847-1072	Sequence 1072, Ap	c 967	28	1.7	1399	9	US-10-174-590-11	Sequence 11, Appl
c 895	28	1.7	577	9	US-09-796-692-4683	Sequence 4683, Ap	c 968	28	1.7	1399	9	US-10-176-758-11	Sequence 11, Appl

969	28	1.7	1399	9	US-10-175-737-11	Sequence 11, Appl
970	28	1.7	1399	9	US-09-993-667-27	Sequence 27, Appl
971	28	1.7	1399	9	US-10-173-706-11	Sequence 11, Appl
972	28	1.7	1399	9	US-10-175-738-11	Sequence 11, Appl
973	28	1.7	1399	9	US-10-175-752-11	Sequence 11, Appl
974	28	1.7	1399	9	US-10-176-482-11	Sequence 11, Appl
975	28	1.7	1399	9	US-10-176-482-11	Sequence 11, Appl
976	28	1.7	1399	9	US-10-176-757-11	Sequence 11, Appl
977	28	1.7	1399	9	US-10-176-913-11	Sequence 11, Appl
978	28	1.7	1399	9	US-10-180-552-11	Sequence 11, Appl
979	28	1.7	1399	9	US-10-180-557-11	Sequence 11, Appl
980	28	1.7	1399	9	US-09-990-438-27	Sequence 27, Appl
981	28	1.7	1399	9	US-09-990-562-27	Sequence 27, Appl
982	28	1.7	1399	9	US-09-997-428-27	Sequence 27, Appl
983	28	1.7	1399	9	US-09-997-666-27	Sequence 27, Appl
984	28	1.7	1399	9	US-10-173-700-11	Sequence 11, Appl
985	28	1.7	1399	9	US-10-174-572-11	Sequence 11, Appl
986	28	1.7	1399	9	US-10-174-579-11	Sequence 11, Appl
987	28	1.7	1399	9	US-10-174-582-11	Sequence 11, Appl
988	28	1.7	1399	9	US-10-174-588-11	Sequence 11, Appl
989	28	1.7	1399	9	US-10-175-739-11	Sequence 11, Appl
990	28	1.7	1399	9	US-10-175-740-11	Sequence 11, Appl
991	28	1.7	1399	9	US-10-175-743-11	Sequence 11, Appl
992	28	1.7	1399	9	US-10-176-488-11	Sequence 11, Appl
993	28	1.7	1399	9	US-10-176-492-11	Sequence 11, Appl
994	28	1.7	1399	9	US-10-176-747-11	Sequence 11, Appl
995	28	1.7	1399	9	US-10-176-750-11	Sequence 11, Appl
996	28	1.7	1399	9	US-10-176-985-11	Sequence 11, Appl
997	28	1.7	1399	9	US-10-176-987-11	Sequence 11, Appl
998	28	1.7	1399	9	US-10-176-991-11	Sequence 11, Appl
999	28	1.7	1399	9	US-10-176-992-11	Sequence 11, Appl
1000	28	1.7	1399	9	US-10-176-993-11	Sequence 11, Appl
	28	1.7	1399	9	US-10-184-658-11	Sequence 11, Appl

## ALIGNMENTS

```

RESULT 1
US-09-764-877-3251/c
; Sequence 3251, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764.877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3251
; LENGTH: 32177
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3251

Query Match          2.9%   Score 49;   DB 10;   Length 32177;
Best Local Similarity 100.0%;   Pred. No. 1.6e-08;
Matches 49;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY  138  CAGGAGTTCAGACTAGCCTGCCCAACATGCTGAAACCCCTATCTCTACT 186
      |||||||
Db    3375  CAGGAGTTCAGACTAGCCTGCCCAACATGCTGAAACCCCTATCTCTACT 3327

RESULT 2
US-09-859-888-3/c
; Sequence 3, Application US/09859888
; Patent No. US20020173459A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND

```

```
; Publication No. US20020192695A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: PIBS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-074C
; CURRENT APPLICATION NUMBER: US/10/161,510
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 133893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-161-510-1

Query Match      2.7%; Score 46; DB 9; Length 133893;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCACGCTACTTAGGAGGCTGAGGCAGGAGGAATTGCTTG 265
|||||
Db 126262 CCTGTAATCCACGCTACTTAGGAGGCTGAGGCAGGAGGAATTGCTTG 126217

RESULT 5
US-09-982-091A-5/c
; Sequence 5, Application US/09982091A
; Patent No. US20020151030A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: KUMAGAI, AKIHO
; APPLICANT: DUNPHY, WILLIAM
; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
; FILE REFERENCE: CITI320-1
; CURRENT APPLICATION NUMBER: US/09/982,091A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US 60/241,246
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 58837
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-982-091A-5

Query Match      2.6%; Score 43; DB 10; Length 58837;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 GGTGGCTCATGCCTATAATCCCGACACTTTGGGAGCCTGAGGT 119
|||||
Db 44531 GGTGGCTCATGCCTATAATCCCGACACTTTGGGAGCCTGAGGT 44489

RESULT 6
US-09-867-701-8121/c
; Sequence 8121, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7324
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-701-8121
```

```
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8121
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(397)
; OTHER INFORMATION: n u A,T,C or G
US-09-867-701-8121

Query Match      2.5%; Score 42; DB 10; Length 397;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 CACCTGTAAATCCAGCTACTTAGGAGGCTGAGGCAGGAGGAAT 259
|||||
Db 139 CACCTGTAAATCCAGCTACTTAGGAGGCTGAGGCAGGAGGAAT 98

RESULT 7
US-09-962-832-154/c
; Sequence 154, Application US/09962832
; Patent No. US20020110821A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; TITLE OF INVENTION: Sels
; FILE REFERENCE: 689290-74
; CURRENT APPLICATION NUMBER: US/09/962,832
; CURRENT FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,077
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,280
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 259
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 154
; LENGTH: 302250
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-832-154

Query Match      2.5%; Score 42; DB 10; Length 302250;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCTGGCCCAACATGTTGAAACCCCT 177
|||||
Db 156903 GTCAGGAGTTCAAGACTAGCTGGCCCAACATGTTGAAACCCCT 156862

RESULT 8
US-09-867-701-7324
; Sequence 7324, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7324
; LENGTH: 400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-701-7324
```



US-09-867-701-7324

Query Match 2.4%; Score 41; DB 10; Length 400;  
Best Local Similarity 100.0%; Pred. No. 2e-05;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAACCC 176  
|||||  
Db 175 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAACCC 215

RESULT 9

US-09-880-107-3872  
; Sequence 3872, Application US/09880107  
; Patent No. US20020142981A1  
; GENERAL INFORMATION:  
; APPLICANT: Horne, Darci T.  
; APPLICANT: Vockley, Joseph G.  
; APPLICANT: Scherf, Uwe  
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
; FILE REFERENCE: 44921-5028-WO  
; CURRENT APPLICATION NUMBER: US/09/880,107  
; CURRENT FILING DATE: 2001-06-14  
; PRIOR APPLICATION NUMBER: US 60/211,379  
; PRIOR FILING DATE: 2000-06-14  
; PRIOR APPLICATION NUMBER: US 60/237,054  
; PRIOR FILING DATE: 2000-10-02  
; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3872  
; LENGTH: 2694  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 Z28339  
US-09-880-107-3872

Query Match 2.4%; Score 41; DB 10; Length 2694;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GGGTGCGGTGCTCATGCCCTATATCCAGCACCTTTGGGAG 111  
|||||  
Db 2130 GGGTGGGTGCTCATGCCCTATATCCAGCACCTTTGGGAG 2170

RESULT 10

US-09-954-456-45  
; Sequence 45, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand  
; TITLE OF INVENTION: Sets  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720

; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 45  
; LENGTH: 3088  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-45

Query Match 2.4%; Score 41; DB 10; Length 3088;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAACCC 176  
|||||  
Db 1364 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAACCC 1404

RESULT 11

US-09-954-456-1621  
; Sequence 1621, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C  
; TITLE OF INVENTION: Sets  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 1621  
; LENGTH: 3088  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-1621

Query Match 2.4%; Score 41; DB 10; Length 3088;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAACCC 176  
|||||  
Db 1364 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAACCC 1404

RESULT 12

US-09-969-347-234  
; Sequence 234, Application US/09969347



```
; Patent No. US20020115085A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: Sets
; CURRENT APPLICATION NUMBER: 689290-69
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/60/237,598
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,604
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 318
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 234
; LENGTH: 3088
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-347-234

Query Match      2.4%; Score 41; DB 10; Length 3088;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGGAACCC 176
      |||||||
Db 1364 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGGAACCC 1404

RESULT 13
US-09-764-860-1091/c
; Sequence 1091, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1091
; LENGTH: 15500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-1091

Query Match      2.4%; Score 41; DB 10; Length 15500;
Best Local Similarity 100.0%; Pred. No. 9.4e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGGAACCC 176
      |||||||
Db 7028 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGGAACCC 5988

RESULT 14
US-09-764-855-321/c
; Sequence 321, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL10
; CURRENT APPLICATION NUMBER: US/09/764,855
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 321
; LENGTH: 16552
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-764-855-321

Query Match      2.4%; Score 41; DB 10; Length 16552;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGGAACCC 176
      |||||||
Db 6659 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGGAACCC 6619

RESULT 15
US-09-764-855-322/c
; Sequence 322, Application US/09764855
; Patent No. US20020119919A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL10
; CURRENT APPLICATION NUMBER: US/09/764,855
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 322
; LENGTH: 16552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-855-322

Query Match      2.4%; Score 41; DB 10; Length 16552;
Best Local Similarity 100.0%; Pred. No. 9.3e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGGAACCC 176
      |||||||
Db 6659 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGGAACCC 6619

RESULT 16
US-09-764-877-3806/c
; Sequence 3806, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3806
; LENGTH: 18878
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3806

Query Match      2.4%; Score 41; DB 10; Length 18878;
Best Local Similarity 100.0%; Pred. No. 9e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 227 TCCACGCTACTTAGGAGGCTGAGCGCAGGAGAAATTGCTTGAA 267
      |||||||
Db 1561 TCCACGCTACTTAGGAGGCTGAGCGCAGGAGAAATTGCTTGAA 1521

RESULT 17
US-09-835-081-3
; Sequence 3, Application US/09835081
; Patent No. US20020151020A1
; GENERAL INFORMATION:
```

```
; APPLICANT: YAN, Xianghe et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001224
; CURRENT APPLICATION NUMBER: US/09/835,081
; CURRENT FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 57130
; TYPE: DNA
; ORGANISM: Human
; US-09-835-081-3

Query Match      2.4%; Score 41; DB 10; Length 57130;
Best Local Similarity 100.0%; Pred. No. 7.1e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GGGTGGGTGGCTCATGCTATATATCCAGCAGCTTTGGGAG 111
|||||
Db 23473 GGGTGGGTGGCTCATGCTATATATCCAGCAGCTTTGGGAG 23513

RESULT 18
US-09-982-091A-5
; Sequence 5, Application US/09982091A
; Patent No. US20020151030A1
; GENERAL INFORMATION:
; APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY
; APPLICANT: KUMAGAI, Akiko
; APPLICANT: DUNPHY, William
; TITLE OF INVENTION: CLASPIN PROTEINS AND METHODS OF USE THEREOF
; FILE REFERENCE: CIT1320-1
; CURRENT APPLICATION NUMBER: US/09/982,091A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US 60/241,246
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 58837
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-982-091A-5

Query Match      2.4%; Score 41; DB 10; Length 58837;
Best Local Similarity 100.0%; Pred. No. 7.1e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCTGCGCCCAACATGGTGAAACCC 176
|||||
Db 1206 GTCAGGAGTTCAAGACTAGCTGCGCCCAACATGGTGAAACCC 1246

RESULT 19
US-09-901-152-3/c
; Sequence 3, Application US/09901152
; Publication No. US20030022824A1
; GENERAL INFORMATION:
; APPLICANT: HU, Song et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001248
; CURRENT APPLICATION NUMBER: US/09/901,152
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 58985
; TYPE: DNA
; ORGANISM: Human
; US-09-901-152-3
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(58985)
; OTHER INFORMATION: n = A,T,C or G
; US-09-901-152-3

Query Match      2.4%; Score 41; DB 9; Length 58985;
Best Local Similarity 100.0%; Pred. No. 7.1e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GGGTGGGTGGCTCATGCTATATATCCAGCAGCTTTGGGAG 111
|||||
Db 15004 GGGTGGGTGGCTCATGCTATATATCCAGCAGCTTTGGGAG 14964

RESULT 20
US-09-859-888-3
; Sequence 3, Application US/09859888
; Patent No. US20020173459A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL001239
; CURRENT APPLICATION NUMBER: US/09/859,888
; CURRENT FILING DATE: 2001-05-18
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 65464
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(65464)
; OTHER INFORMATION: n = A,T,C or G
; US-09-859-888-3

Query Match      2.4%; Score 41; DB 9; Length 65464;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 CACCTGTATATCCAGCTACTTAGGAGGCTGAGCAGGAGAA 258
|||||
Db 7275 CACCTGTATATCCAGCTACTTAGGAGGCTGAGCAGGAGAA 7315

RESULT 21
US-09-880-107-3949/c
; Sequence 3949, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3949
; LENGTH: 76798
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 284718
```

US-09-880-107-3949

Query Match 2.4%; Score 41; DB 10; Length 76798;  
Best Local Similarity 100.0%; Pred. No. 6.7e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GGGTGGGTGGCTGATCCATATCCAGCACTTTGGGAG 111  
|||||  
Db 28999 GGGTGGGTGGCTGATCCATATATCCAGCACTTTGGGAG 28959

RESULT 22

US-10-094-989-3  
; Sequence 3, Application US/10094989  
; Patent No. US20020115179A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE  
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN  
; TITLE OF INVENTION: PHOSPHODIESTERASE PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL001063DIV  
; CURRENT APPLICATION NUMBER: US/10/094,989  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: 09/754,250  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 111282  
; TYPE: DNA  
; ORGANISM: Homo saplen  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(111282)  
; OTHER INFORMATION: n = A,T,C or G

US-10-094-989-3

Query Match 2.4%; Score 41; DB 12; Length 111282;  
Best Local Similarity 100.0%; Pred. No. 6.2e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 TCCAGCTACTTAGGAGCTGAGCAGGAGAATTGCTTGAA 267  
|||||  
Db 3839 TCCAGCTACTTAGGAGCTGAGCAGGAGAATTGCTTGAA 3879

RESULT 23

US-09-804-474A-3/c  
; Sequence 3, Application US/09804474A  
; Patent No. US20020119518A1  
; GENERAL INFORMATION:  
; APPLICANT: KODET, Stefan et al  
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CL000891  
; CURRENT APPLICATION NUMBER: US/09/804,474A  
; CURRENT FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 126512  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(126512)  
; OTHER INFORMATION: n = A,T,C or G

US-09-804-474A-3

Query Match 2.4%; Score 41; DB 10; Length 126512;  
Best Local Similarity 100.0%; Pred. No. 6e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 227 TCCAGCTACTTAGGAGCTGAGCAGGAGAATTGCTTGAA 267  
|||||  
Db 10106 TCCAGCTACTTAGGAGCTGAGCAGGAGAATTGCTTGAA 10066

RESULT 24

US-09-969-708-79/c  
; Sequence 79, Application US/09969708  
; Patent No. US20020102532A1  
; GENERAL INFORMATION:  
; APPLICANT: Augustus, Meena  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign  
; TITLE OF INVENTION: Sets  
; FILE REFERENCE: 689290-70  
; CURRENT APPLICATION NUMBER: US/09/969,708  
; CURRENT FILING DATE: 2001-10-03  
; PRIOR APPLICATION NUMBER: US/60/237,606  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: US/60/237,608  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: US/60/237,425  
; PRIOR FILING DATE: 2000-10-03  
; NUMBER OF SEQ ID NOS: 658  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 79  
; LENGTH: 145831  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-969-708-79

Query Match 2.4%; Score 41; DB 10; Length 145831;  
Best Local Similarity 100.0%; Pred. No. 5.9e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAAACCC 176  
|||||  
Db 113813 GTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAAACCC 113773

RESULT 25

US-09-954-456-2116/c  
; Sequence 2116, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C  
; TITLE OF INVENTION: Sets  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0

```

; SEQ ID NO 2116
; LENGTH: 145831
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-2116

Query Match      2.4%; Score 41; DB 10; Length 145831;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCTGGCCACATGGTGAACCC 176
|||||
Db 113813 GTCAGGAGTTCAAGACTAGCTGGCCACATGGTGAACCC 113773

RESULT 26
US-09-880-107-832/c
; Sequence 832, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 832
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA401958
US-09-880-107-832

Query Match      2.4%; Score 40; DB 10; Length 470;
Best Local Similarity 100.0%; Pred. No. 4.3e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTATCCAGCTACTTAGGAGGCTGAGGAGGAGGAGAT 259
|||||
Db 127 CCTGTATCCAGCTACTTAGGAGGCTGAGGAGGAGGAGAT 88

RESULT 27
US-09-729-674-51/c
; Sequence 51, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallie, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Agostino, Michael J.
; APPLICANT: Steindinger II, Robert J.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: 6055-64X
; CURRENT APPLICATION NUMBER: US/09/729,674

```

```

; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/539,330
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 283
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-729-674-51

Query Match      2.4%; Score 40; DB 10; Length 2791;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTATCCAGCTACTTAGGAGGCTGAGGAGGAGGAGAT 259
|||||
Db 246 CCTGTATCCAGCTACTTAGGAGGCTGAGGAGGAGGAGAT 207

RESULT 28
US-09-764-847-1487/c
; Sequence 1487, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1487
; LENGTH: 5815
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1487

Query Match      2.4%; Score 40; DB 10; Length 5815;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTATCCAGCTACTTAGGAGGCTGAGGAGGAGGAGAT 259
|||||
Db 2561 CCTGTATCCAGCTACTTAGGAGGCTGAGGAGGAGGAGAT 2522

RESULT 29
US-09-764-853-896/c
; Sequence 896, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 896
; LENGTH: 7373
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-853-896

Query Match      2.4%; Score 40; DB 10; Length 7373;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 GGTGCGGTGGCTCATCGCTATTAATCCACGACTTTGGGAG 111
|||||

```

```
Db 2095 GGTCCGGTGGCTCATGCTATATATCCAGCAGCTTTGGGAG 2056
; LENGTH: 13069
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1850

Query Match 2.4%; Score 40; DB 10; Length 13059;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 30
US-09-764-877-2718
; Sequence 2718, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2718
; LENGTH: 9968
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-2718

Query Match 2.4%; Score 40; DB 10; Length 9968;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 TCAGGAGTTCAAGACTAGCCTGGCCACATGGTGAACCC 176
|||||
Db 1518 TCAGGAGTTCAAGACTAGCCTGGCCACATGGTGAACCC 1557

RESULT 31
US-09-764-868-1471/c
; Sequence 1471, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1471
; LENGTH: 10322
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-868-1471

Query Match 2.4%; Score 40; DB 9; Length 10322;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 CAGAGTTCAAGACTAGCCTGGCCACATGGTGAACCC 177
|||||
Db 10029 CAGAGTTCAAGACTAGCCTGGCCACATGGTGAACCC 9990

RESULT 32
US-09-764-869-1850/c
; Sequence 1850, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1850
```

```
; LENGTH: 13069
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1850

Query Match 2.4%; Score 40; DB 10; Length 13059;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAT 259
|||||
Db 6200 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAT 6161

RESULT 33
US-09-816-248-18
; Sequence 18, Application US/09816248
; Patent No. US20020137703A1
; GENERAL INFORMATION:
; APPLICANT: BAUMANN, PETER
; APPLICANT: CECH, THOMAS R.
; TITLE OF INVENTION: PROTECTION-OF-TELOMERE-1 (POT-1) PROTEIN AND ENCODING
; FILE REFERENCE: POLYNUCLEOTIDES
; CURRENT APPLICATION NUMBER: US/09/816,248
; CURRENT FILING DATE: 2001-03-26
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 27377
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-816-248-18

Query Match 2.4%; Score 40; DB 10; Length 27377;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAT 259
|||||
Db 5540 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAT 5579

RESULT 34
US-09-817-198A-3/c
; Sequence 3, Application US/09817198A
; Patent No. US20020146758A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001188
; CURRENT APPLICATION NUMBER: US/09/817,198A
; CURRENT FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 28770
; TYPE: DNA
; ORGANISM: Human
US-09-817-198A-3

Query Match 2.4%; Score 40; DB 10; Length 28770;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAT 259
|||||
Db 7724 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAT 7685

RESULT 35
```

```
US-09-764-869-1849/c
; Sequence 1849, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1849
; LENGTH: 32203
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1849

Query Match          2.4%; Score 40; DB 10; Length 32203;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAAT 259
|||||
Db 6194 CCTGTAATCCAGCTACTTACTAGGAGGCTGAGGCAGGAGAAAT 6155

RESULT 36
US-09-918-686-2/c
; Sequence 2, Application US/09918686
; Patent No. US20020076720A1
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Prohl, Sean
; APPLICANT: Paepfer, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 51719
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1246, 2572, 2604
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-2

Query Match          2.4%; Score 40; DB 10; Length 51719;
Best Local Similarity 100.0%; Pred. No. 1.6e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAAT 259
|||||
Db 25477 CCTGTAATCCAGCTACTTACTAGGAGGCTGAGGCAGGAGAAAT 25438

RESULT 37
US-09-962-436-36
; Sequence 36, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; CURRENT FILING DATE: 2001-09-25
```

```
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36
; LENGTH: 84539
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-36

Query Match          2.4%; Score 40; DB 10; Length 84539;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAAT 259
|||||
Db 6846 CCTGTAATCCAGCTACTTACTAGGAGGCTGAGGCAGGAGAAAT 6885

RESULT 38
US-09-918-686-1/c
; Sequence 1, Application US/09918686
; Patent No. US20020076720A1
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Prohl, Sean
; APPLICANT: Paepfer, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 92139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7043, 8369, 8401
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

Query Match          2.4%; Score 40; DB 10; Length 92139;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGAAAT 259
|||||
Db 31274 CCTGTAATCCAGCTACTTACTAGGAGGCTGAGGCAGGAGAAAT 31235

RESULT 39
US-09-880-107-1542/c
; Sequence 1542, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
```

; NUMBER OF SEQ ID NOS: 3950  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1542  
; LENGTH: 110096  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AD0000092  
US-09-880-107-1542

Query Match 2.4%; Score 40; DB 10; Length 110096;  
Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGCGAGGAAT 259  
|||||  
Db 70191 CCTGTAATCCAGCTACTTAGGAGGCTGAGCGAGGAAT 70152

## RESULT 40

; Sequence 2735, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2735  
; LENGTH: 349  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-2735

Query Match 2.3%; Score 39; DB 10; Length 349;  
Best Local Similarity 100.0%; Pred. No. 9.9e-05;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAAC 174  
|||||  
Db 197 GTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAAC 235

## RESULT 41

; Sequence 2736, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2736  
; LENGTH: 425  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-2736

Query Match 2.3%; Score 39; DB 10; Length 425;  
Best Local Similarity 100.0%; Pred. No. 9.5e-05;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAAC 174  
|||||  
Db 197 GTCAGGAGTTCAAGACTAGCTGGCCCAACATGGTGAAC 235

## RESULT 42

US-09-822-830A-569/c  
; Sequence 569, Application US/09822830A  
; Patent No. US20020142952A1  
; GENERAL INFORMATION:  
; APPLICANT: Genetics Institute, Inc.  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Resnick, Richard J.  
; APPLICANT: Gulukota, Kamalakar  
; APPLICANT: Graham, James R.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
; FILE REFERENCE: GIN 6402  
; CURRENT APPLICATION NUMBER: US/09/822,830A  
; CURRENT FILING DATE: 2001-03-29  
; PRIOR APPLICATION NUMBER: 60/195,604  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 631  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 569  
; LENGTH: 1788  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-822-830A-569

Query Match 2.3%; Score 39; DB 10; Length 1788;  
Best Local Similarity 100.0%; Pred. No. 7e-05;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 TAATCCAGCACCTTTGGAGCCTGAGGTGGGTGGATCAC 130  
|||||  
Db 706 TAATCCAGCACCTTTGGAGCCTGAGGTGGGTGGATCAC 668

## RESULT 43

US-09-764-860-1106/c  
; Sequence 1106, Application US/09764860  
; Patent No. US20020094953A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC008  
; CURRENT APPLICATION NUMBER: US/09/764,860  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1198  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1106  
; LENGTH: 8082  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-860-1106

Query Match 2.3%; Score 39; DB 10; Length 8082;  
Best Local Similarity 100.0%; Pred. No. 5.1e-05;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 CTGTAATCCAGCTACTTAGGAGGCTGAGCGAGGAAT 259  
|||||  
Db 3411 CTGTAATCCAGCTACTTAGGAGGCTGAGCGAGGAAT 3373

## RESULT 44

US-09-764-869-2209/c  
; Sequence 2209, Application US/09764869  
; Patent No. US20020061521A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC007  
; CURRENT APPLICATION NUMBER: US/09/764,869  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 2442  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2209  
; LENGTH: 32190  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-869-2209

Query Match 2.3%; Score 39; DB 10; Length 32190;  
Best Local Similarity 100.0%; Pred. No. 3.8e-05;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 TAATCCAGCAGCTTTGGAGGCTGAGGTGGTGGATCAC 130  
|||||  
Db 23758 TAATCCAGCAGCTTTGGAGGCTGAGGTGGTGGATCAC 23720

## RESULT 45

US-09-764-869-1556  
; Sequence 1556, Application US/09764869  
; Patent No. US20020061521A1  
; GENERAL INFORMATION:

; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC007  
; CURRENT APPLICATION NUMBER: US/09/764,869  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 2442  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1556  
; LENGTH: 26048  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-869-1556

Query Match 2.3%; Score 38; DB 10; Length 26048;  
Best Local Similarity 100.0%; Pred. No. 8.7e-05;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 CACCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGA 255  
|||||  
Db 18573 CACCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGA 18610

## RESULT 46

US-09-967-768A-316/c  
; Sequence 316, Application US/09967768A  
; Patent No. US20020150877A1  
; GENERAL INFORMATION:

; APPLICANT: Augustus, Meena  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu  
; FILE REFERENCE: 689290-72  
; CURRENT APPLICATION NUMBER: US/09/967,768A  
; CURRENT FILING DATE: 2001-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,109  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,034  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,111  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 325  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 316  
; LENGTH: 143068  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-967-768A-316

Query Match 2.3%; Score 38; DB 10; Length 143068;  
Best Local Similarity 100.0%; Pred. No. 6.1e-05;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 GTGCGGTGGCTCATGCTATATATCCAGCAGCTTTGGGA 110  
|||||  
Db 56454 GTGCGGTGGCTCATGCTATATATCCAGCAGCTTTGGGA 56417

## RESULT 47

US-09-962-436-281  
; Sequence 281, Application US/09962436  
; Patent No. US20020081301A1  
; GENERAL INFORMATION:

; APPLICANT: Soppet, Daniel  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign  
; FILE REFERENCE: 689290-75  
; CURRENT APPLICATION NUMBER: US/09/962,436  
; CURRENT FILING DATE: 2001-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,082  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/234,924  
; PRIOR FILING DATE: 2000-09-25  
; NUMBER OF SEQ ID NOS: 568  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 281  
; LENGTH: 167343  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-962-436-281

Query Match 2.3%; Score 38; DB 10; Length 167343;  
Best Local Similarity 100.0%; Pred. No. 5.9e-05;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGA 257  
|||||  
Db 4607 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGA 4644

## RESULT 48

US-09-964-824A-273  
; Sequence 273, Application US/09964824A  
; Patent No. US20020102531A1  
; GENERAL INFORMATION:

; APPLICANT: Horrigan, Stephen  
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign  
; FILE REFERENCE: 689290-73  
; CURRENT APPLICATION NUMBER: US/09/964,824A  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US/60/236,033  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,032  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US/60/236,028  
; PRIOR FILING DATE: 2000-09-28  
; NUMBER OF SEQ ID NOS: 583  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 273  
; LENGTH: 167343  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-964-824A-273

Query Match 2.3%; Score 38; DB 10; Length 167343;  
Best Local Similarity 100.0%; Pred. No. 5.9e-05;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CCTGTAATCCAGCTACTTAGGAGGCTGAGGCAGGAGA 257



Db 4607 CCTGTAATCCCGACTACTTAGGAGGCTGAGGCAGGAGA 4644

RESULT 49  
US-09-764-877-376  
; Sequence 376, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 376  
; LENGTH: 293  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (287)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-877-376

Query Match 2.2%; Score 37; DB 10; Length 293;  
Best Local Similarity 100.0%; Pred. No. 0.00049;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAA 172  
Db 199 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAA 235

RESULT 50  
US-09-867-701-5962  
; Sequence 5962, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5962  
; LENGTH: 438  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-5962

Query Match 2.2%; Score 37; DB 10; Length 438;  
Best Local Similarity 100.0%; Pred. No. 0.00045;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAA 172  
Db 345 GTCAGGAGTTCAAGACTAGCCTGGCCCAACATGGTGAA 381

Search completed: February 12, 2003, 16:23:33  
Job time : 1926 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:51:09 ; Search time 147.133 Seconds  
(without alignments)  
5613.764 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_5\_55

Perfect score: 51  
Sequence: 1 tcaagaccagcctggacac.....sggtctotacaaaaataca 51

Scoring table: IDENTITY\_NUC  
Gapop 10\_0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_nam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	96.1	364	17	AQ428360 CITBI-E1-
2	49	96.1	449	17	AQ424894 CITBI-E1-
3	44.2	86.7	194	12	BF955584 CMO-NN024
4	44.2	86.7	227	14	N74947 za45g05.s1
5	44.2	86.7	465	17	AQ006763 CIT-HSP-2
6	43	84.3	422	17	B76188 RPC11-1301

7	43	84.3	646	17	AG123613
8	42.6	83.5	147	17	AQ394280 CITBI-E1-
9	42.6	83.5	149	10	AW880489 QVO-OT003
10	42.6	83.5	159	10	AW846743 QV3-CT019
11	42.6	83.5	227	9	AA595144 n184a11.s
12	42.6	83.5	246	12	BF819673 MRI-RT002
13	42.6	83.5	268	9	AI970561 wr12u09.x
14	42.6	83.5	274	13	BG990678 RC4-HT109
15	42.6	83.5	303	10	AW103010 xd41h11.x
16	42.6	83.5	310	9	AA326336 EST29495
17	42.6	83.5	369	9	AA484273 nc88u08.s
18	42.6	83.5	384	9	AA041276 z108u03.s
19	42.6	83.5	406	9	AA484304 nc89u08.s
20	42.6	83.5	418	17	BH610078 HIV20A05
21	42.6	83.5	439	9	AI910067 RC-BT234-
22	42.6	83.5	447	9	AI811345 lw38f06.x
23	42.6	83.5	484	9	AL699717 DKFZP686C
24	42.6	83.5	511	17	A2518820 RPC1-11-3
25	42.6	83.5	541	17	AQ897411 HS-3153_A
26	42.6	83.5	571	9	AA131088 zol16c04.s
27	42.6	83.5	588	17	AQ581201 RPC1-11-4
28	42.6	83.5	634	9	AI561147 Lq37u05.x
29	42.6	83.5	645	9	AI065029 HA0859 HU
30	42.6	83.5	650	17	AG035121 Pan croql
31	42.6	83.5	690	17	AQ415537 RPC1-11-1
32	42.6	83.5	748	9	AL700218 DKFZP686N
33	42.6	83.5	759	12	BG434724 602507926
34	42.6	83.5	832	17	AQ752230 HS-5572.B
35	42.6	83.5	947	12	BF794959 602256853
36	42.6	83.5	1131	14	BM799755 AGENCOURT
37	42.6	83.5	1254	13	BM548499 AGENCOURT
38	41.6	81.6	172	17	AQ756322 HS-5201_B
39	41.6	81.6	489	17	AQ433028 HS-5148.B
40	41.6	81.6	497	13	BI491471 dfl10a02.w
41	41.6	81.6	514	10	AW020452 dfl10a02.y
42	41.6	81.6	624	17	AQ783398 HS-5563_A
43	41.4	81.2	505	17	B46917 HS-1066-A2-
44	41.4	81.2	583	10	AV718985 AV718985
45	41.4	81.2	590	10	AV720747 AV720747

ALIGNMENTS

RESULT 1	AQ428360	CITBI-EI-2576K5.TF	364 bp	DNA	linear	GSS 24-MAR-1999
LOCUS	AQ428360	CITBI-EI Homo sapiens	genomic clone	2576K5	DNA	
DEFINITION	AQ428360	sequence.				
ACCESSION	AQ428360	GI:4496126				
VERSION	AQ428360					
KEYWORDS	GSS					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 364)					
AUTHORS	Zhao, S., Adams, M.D., Niernan, W., Malek, J., Shizuya, H., Simon, M. and Venter, J.C.					
TITLE	Use of BAC End Sequences from CalTech Libraries for Sequence-Ready Map Building					
JOURNAL	Unpublished (1997)					
COMMENT	Contact: Shaying Zhao, William Niernan, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html. Seq primer: M13-21					

```

FEATURES
  source
    Class: BAC ends.
    Location/Qualifiers
      1..364
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="2576K5"
        /clone_lib="CITBI-El"
        /sex="male"
        /cell_type="sperm"
        /note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
          CalTech Human BAC Library D"
BASE COUNT      106 a      82 c      98 g      78 t
ORIGIN
Query Match      96.1%; Score 49; DB 17; Length 364;
Best Local Similarity 96.1%; Pred. No. 8.1e-06;
Matches 49; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGACAACCTTGGAGAACCCGGTCTCTACAAAAATACA 51
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 TCAAGACCGCTGTACAACTTGGAGAACCCGGTCTCTACAAAAATACA 52
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
A0424894
LOCUS
DEFINITION      A0424894 449 bp DNA linear GSS 24-MAR-1999
COMMENT         CITBI-El-2576P11.TF CITBI-El Homo sapiens genomic clone 2576P11,
                  DNA sequence.
ACCESSION      A0424894
VERSION        A0424894.1 GI:4498160
KEYWORDS       GSS.
SOURCE         human.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 449)
AUTHORS        Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
                Venter,J.C.
TITLE          Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
                Map Building
JOURNAL        Unpublished (1997)
COMMENT        Other_GSSs: CITBI-El-2576P11.TR
                Contact: Shaying Zhao, William Nierman, Mark Adams
                Department of Eukaryotic Genomics
                The Institute for Genomic Research
                9712 Medical Center Dr., Rockville, MD 20850
                Tel: 301 838 0200
                Fax: 301 838 0208
                Email: hbeatigr.org
                Clones are available from Research Genetics (info@resgen.com). BAC
                end search page:
                http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
                Seq primer: M13-21
                Class: BAC ends.
FEATURES
  source
    Location/Qualifiers
      1..449
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="2576P11"
        /clone_lib="CITBI-El"
        /sex="male"
        /cell_type="sperm"
        /note="Vector: pBelOBAC11; Site_1: EcoRI; Site_2: EcoRI;
          Caltech Human BAC Library D"
BASE COUNT      134 a      98 c      123 g      94 t
ORIGIN
Query Match      96.1%; Score 49; DB 17; Length 449;
Best Local Similarity 96.1%; Pred. No. 8.2e-06;
Matches 49; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGACAACCTTGGAGAACCCGGTCTCTACAAAAATACA 51
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 2 TCAAGACCGCTGTACAACTTGGAGAACCCGGTCTCTACAAAAATACA 52
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 3
BF955584
LOCUS
DEFINITION      BF955584 194 bp mRNA linear EST 22-JAN-2001
COMMENT         CM0-NN0249-181100-710-e09 NN0249 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BF955584
VERSION        BF955584.1 GI:12372859
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 194)
AUTHORS        Dias Neto,F., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
                Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
                Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
                Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
                ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
                Simpson,A.J.
TITLE          Shotgun sequencing of the human transcriptome with ORF expressed
                sequence tags
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE        20202663
COMMENT        Contact: Simpson A.J.G.
                Laboratory of Cancer Genetics
                Ludwig Institute for Cancer Research
                Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                Brazil
                Tel: +55-11-2704922
                Fax: +55-11-2707001
                Email: asimpson@ludwig.org.br
                This sequence was derived from the FAPESP/LICR Human Cancer Genome
                Project. This entry can be seen in the following URL
                (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM0&t2=CM0-NN0249-
                181100-710-e09&t3=2000-11-18&t4=1)
                Seq primer: puc 18 forward
                High quality sequence stop: 194.
FEATURES
  source
    Location/Qualifiers
      1..194
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_lib="NN0249"
        /dev_stage="Adult"
        /note="Organ: nervous normal; Vector: puc18; Site_1: SmaI;
          Site_2: SmaI; A mini-library was made by cloning products
          derived from ORESTES PCR (U.S. Letters Patent application
          No. 196,716 - Ludwig Institute for Cancer Research)
          profiles into the pUC 18 vector. Reverse transcription of
          tissue mRNA and cDNA amplification were performed under
          low stringency conditions."
BASE COUNT      61 a      48 c      50 g      35 t
ORIGIN
Query Match      86.7%; Score 44.2; DB 12; Length 194;
Best Local Similarity 90.2%; Pred. No. 0.00028;
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGACAACCTTGGAGAACCCGGTCTCTACAAAAATACA 51
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102 TCAAGACCGCTGGACAACATGCGAAAACCCCGTCTCTACAAAAATACA 152
      ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
N74947/c
LOCUS
DEFINITION      N74947 227 bp mRNA linear EST 29-MAR-1996
COMMENT         za45g05.sl Soares fetal liver spleen INFLS Homo sapiens cDNA clone
                IMAGE:295544 3' similar to contains Alu repetitive element;contains
                element PPR7 repetitive element ; mRNA sequence.
ACCESSION      N74947
VERSION        N74947.1 GI:1237493
KEYWORDS       EST.

```

SOURCE  
ORGANISM  
human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 227)  
AUTHORS  
Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman  
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marras,M., Parsons,J.,  
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston  
,R., Williamson,A., Wohldmann,P. and Wilson,R.  
The WashU-Merck EST Project  
TITLE  
JOURNAL  
COMMENT  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL : contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: mi3 -40 forward  
High quality sequence stop: 205.  
FEATURES  
Location/Qualifiers  
1..227  
/organism="Homo sapiens"  
/db\_xref="GDB:1240465"  
/db\_xref="taxon:9606"  
/clone="IMAGE:295544"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="PH10B (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)  
with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACTGGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified p7T3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo."  
BASE COUNT 41 a 60 c 57 g 68 t 1 others  
ORIGIN  
Query Match 86.7%; Score 44.2; DB 14; Length 227;  
Best Local Similarity 90.2%; Pred. No. 0.00028;  
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 TCAAGACCAGCTGGACAACTTGGAGAGACCGTCTCTACAAAAATACA 51  
|||||  
Db 209 TCAAGACCAGCTGGTCAACATGCGAGACCGTCTCTACAAAAATACA 159  
|||||  
RESULT 5  
AQ006763 465 bp DNA linear GSS 27-JUN-1998  
LOCUS CIT-HSP-2291F2.TR CIT-HSP Homo sapiens genomic clone 2291F2, DNA  
DEFINITION sequence.  
ACCESSION AQ006763  
VERSION AQ006763.1 GI:3083989  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 465)  
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuwa,H.,  
Simon,M. and Venter,J.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building (1998)  
TITLE Unpublished (1998)  
JOURNAL Contact: Mark Adams  
COMMENT Department of Eukaryotic Genomics

human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 227)  
AUTHORS Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman  
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marras,M., Parsons,J.,  
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston  
,R., Williamson,A., Wohldmann,P. and Wilson,R.  
The WashU-Merck EST Project  
TITLE JOURNAL COMMENT Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL : contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: mi3 -40 forward  
High quality sequence stop: 205.  
FEATURES Location/Qualifiers  
1..227  
/organism="Homo sapiens"  
/db\_xref="GDB:1240465"  
/db\_xref="taxon:9606"  
/clone="IMAGE:295544"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="PH10B (ampicillin resistant)"  
/note="Organ: Liver and Spleen; Vector: p7T3D (Pharmacia)  
with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACTGGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified p7T3 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M.Fatima Bonaldo."  
BASE COUNT 41 a 60 c 57 g 68 t 1 others  
ORIGIN  
Query Match 86.7%; Score 44.2; DB 14; Length 227;  
Best Local Similarity 90.2%; Pred. No. 0.00028;  
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 TCAAGACCAGCTGGACAACTTGGAGAGACCGTCTCTACAAAAATACA 51  
|||||  
Db 209 TCAAGACCAGCTGGTCAACATGCGAGACCGTCTCTACAAAAATACA 159  
|||||  
RESULT 5  
AQ006763 465 bp DNA linear GSS 27-JUN-1998  
LOCUS CIT-HSP-2291F2.TR CIT-HSP Homo sapiens genomic clone 2291F2, DNA  
DEFINITION sequence.  
ACCESSION AQ006763  
VERSION AQ006763.1 GI:3083989  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 465)  
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuwa,H.,  
Simon,M. and Venter,J.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building (1998)  
TITLE Unpublished (1998)  
JOURNAL Contact: Mark Adams  
COMMENT Department of Eukaryotic Genomics

The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). HAC  
end search page:  
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13 Reverse  
Class: BAC ends.  
FEATURES Location/Qualifiers  
1..465  
/organism="Homo sapiens"  
/db\_xref="GDB:7150997"  
/db\_xref="taxon:9606"  
/clone="2291F2"  
/clone\_lib="CIT-HSP"  
/sex="Male"  
/cell\_type="Sperm"  
/note="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2:  
HindIII"  
BASE COUNT 130 a 105 c 125 g 105 t  
ORIGIN  
Query Match 86.7%; Score 44.2; DB 17; Length 465;  
Best Local Similarity 90.2%; Pred. No. 0.00029;  
Matches 46; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 TCAAGACCAGCTGGACAACTTGGAGAGACCGTCTCTACAAAAATACA 51  
|||||  
Db 301 TCAAGACCAGCTGGACAACTGCGCAACCCGTCTCTACAAAAATACA 351  
|||||  
RESULT 6  
B76188 422 bp DNA linear GSS 08-APR-1999  
LOCUS RPC111-13016.TP RPC1-11 Homo sapiens genomic clone RPC1-11-13016,  
DNA sequence.  
DEFINITION B76188  
ACCESSION B76188  
VERSION B76188.1 GI:2771875  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 422)  
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden  
,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter  
,J.C.  
Use of BAC End Sequences for Sequence-Ready Map Building  
(1997)  
TITLE Unpublished (1997)  
JOURNAL Other\_GSSs: RPC111-13016.TV  
COMMENT Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are derived from the human BAC library RPC1-11. For BAC  
library availability, please contact p.ter de Jong  
(p.terdejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: SP6  
Class: BAC ends.  
FEATURES Location/Qualifiers  
1..422  
/organism="Homo sapiens"  
/db\_xref="GDB:7504959"  
/db\_xref="taxon:9606"  
/clone="RPC1-11-13016"

```

/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT      123 a   96 c  104 g   99 t
ORIGIN

Query Match      84.3%; Score 43; DB 17; Length 422;
Best Local Similarity 90.2%; Pred. No. 0.00069;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACAGCCTGGACAACCTTGGAGAACCSGGTCTCTACAAAAAATACA 51
|||||
Db 220 TCAAGACAGCCTGGACAACATGTGTGAACCTGTCTCTACCAAAAAAATACA 270

RESULT 7
AG123613
LOCUS      646 bp      DNA      linear      GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-133F03.F, genomic survey sequence.
ACCESSION  AG123613
VERSION     AG123613.1 GI:16652778
KEYWORDS   GSS.
SOURCE     Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male
           BAC Library clone:PTB-133F03.F.
ORGANISM   Pan troglodytes
REFERENCE  1
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
           Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of Library PTB
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 646)
AUTHORS    Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
           Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL    Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
           and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
           1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
           (E-mail:chimpesgsc.riken.go.jp, URL:http://hnp.gsc.riken.go.jp/,
           Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT    Clones are derived from the chimpanzee BAC library PTB This BAC end
           was generated during the R&D process and may have higher chance of
           clone tracking errors.
PRIMERS

LIBRARY
Vector      : pKSL45
R.Site 1    : SacI
R.Site 2    : SacI.
Location/Qualifiers
1. 646
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-133F03.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT  200 a  139 c  138 g  167 t  2 others
ORIGIN

Query Match      84.3%; Score 43; DB 17; Length 646;
Best Local Similarity 90.2%; Pred. No. 0.0007;
Matches 46; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACAGCCTGGACAACCTTGGAGAACCSGGTCTCTACAAAAAATACA 51
|||||
Db 99 TCAAGACAGCCTGGACAACATGGCAAAACCTTGTCTCTACAAAAAATACA 149

RESULT 8

```

```

AQ394280
LOCUS      147 bp      DNA      linear      GSS 06-MAR-1999
DEFINITION CITBI-E1-2546J24.TR CITBI-E1 Homo sapiens genomic clone 2546J24,
           DNA sequence.
ACCESSION  AQ394280
VERSION     AQ394280.1 GI:4365303
KEYWORDS   GSS.
SOURCE     human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 147)
AUTHORS    Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
           Venter,J.C.
TITLE      Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
           Map Building
JOURNAL    Unpublished (1997)
COMMENT    Other_GSSs: CITBI-E1-2546J24.TF
           Contact: Shaying Zhao, William Nierman, Mark Adams
           Department of Eukaryotic Genomics
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850
           Tel: 301 838 0200
           Fax: 301 838 0208
           Email: hbest@ig.org
           Clones are available from Research Genetics (info@resgen.com). BAC
           end search page:
           http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
           Seq primer: M13 Reverse
           Class: BAC ends.
FEATURES             Location/Qualifiers
     source            1..147
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="2546J24"
                     /clone_lib="CITBI-E1"
                     /sex="male"
                     /cell_type="sperm"
     note              1. Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
                     Caltech Human BAC Library D
BASE COUNT  45 a   41 c   33 g   28 t
ORIGIN

Query Match      83.5%; Score 42.6; DB 17; Length 147;
Best Local Similarity 88.2%; Pred. No. 0.0009;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACAGCCTGGACAACCTTGGAGAACCSGGTCTCTACAAAAAATACA 51
|||||
Db 21 TCAAGACAGCCTGGCCAACATGGCAAAACCGGTCTCTACTAAAAAATACA 71

RESULT 9
AW880489
LOCUS      149 bp      mRNA      linear      EST 23-MAY-2000
DEFINITION QV0-OT0030-120500-219-g11 OT0030 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW880489
VERSION     AW880489.1 GI:8042499
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 149)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
           Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
           Goldman,G.H., Carvalheiro,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
           Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
           ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
           Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
           sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

```

MEDLINE  
COMMENT

20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-QV0-OT0030-120  
500-219-g1&ts3-2000-05-12&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 6  
High quality sequence stop: 149.  
Location/Qualifiers

FEATURES  
source

1..149  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="OT0030"  
/dev\_stage="Adult"  
/note="Organ: ovary; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (O.S. Letters Patent application No. 196  
,716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 50 a 32 c 42 g 25 t  
ORIGIN

Query Match 83.5%; Score 42.6; DB 10; Length 149;  
Best Local Similarity 88.2%; Pred. No. 0.0009;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCGCTGGACAACCTTGGAGAACCGGCTCTCTACAAAAATACA 51  
|||||  
Db 32 TCAAGACCAGCGCTGGCAACATGGAGAACCGCTCTCTACAAAAATACA 82

RESULT 10  
AW846743/c

LOCUS  
DEFINITION QV3-CT0194-051099-003-cl2 CT0194 Homo sapiens cDNA, mRNA linear EST 19-MAY-2000  
ACCESSION AW846743  
VERSION AW846743.1 GI:7942260  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS

1 (bases 1 to 159)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL  
MEDLINE  
COMMENT

20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL

FEATURES  
source

1..159  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CT0194"  
/dev\_stage="Adult"  
/note="Organ: colon; Vector: puc18; Site\_1: SmaI; Site\_2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (O.S. Letters Patent application No. 196  
,716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

BASE COUNT 36 a 39 c 36 g 48 t  
ORIGIN

Query Match 83.5%; Score 42.6; DB 10; Length 159;  
Best Local Similarity 88.2%; Pred. No. 0.0009;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCGCTGGACAACCTTGGAGAACCGGCTCTCTACAAAAATACA 51  
|||||  
Db 72 TCAAGACCAGCGCTGGCAACATGGCAAAACCGGCTCTCTACAAAAATACA 22

RESULT 11  
AA595144/c

LOCUS  
DEFINITION n184a11.s1 NCI\_CGAP\_Br2 Homo sapiens cDNA clone IMAGE:1057340 3'  
similar to contains Alu repetitive element;contains element MSRI  
repetitive element ;, mRNA sequence.

ACCESSION AA595144  
VERSION AA595144.1 GI:2410494  
KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS

1 (bases 1 to 227)  
NCI\_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL  
COMMENT

Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert length: 1155 Std Error: 0.00  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 219.  
Location/Qualifiers

FEATURES  
source

1..227  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:1057340"  
/clone\_lib="NCI\_CGAP\_Br2"  
/sex="female, f'ooled"  
/tissue\_type="breast"  
/lab\_host="DH10B"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from pooled bulk  
oligo tumor tissue, and was then primed with a Not I -  
oligo(dT) primer. Double-stranded cDNA was ligated to Eco

```

RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7n3
vector. This library is the normalized version of
NCI CGAP Brl.1. Library was constructed by Bento Soares
and M. Fatima Bonaldo.
BASE COUNT      43 a   68 c   51 g   65 t
ORIGIN
Query Match      83.5%;   Score 42.6;   DB 9;   Length 227;
Best Local Similarity 88.2%;   Pred. No. 0.00091;
Matches 45;   Conservative 1;   Mismatches 5;   Indels 0;   Gaps 0;

Qy 1 TCAAGACAGCCTGGACAACTTGGAGAACCGGTCTCTCTACAAAAATACA 51
|||||
Db 183 TCAAGACAGCCTGGGCAACATGGAGAACCCCTGCTCTCTACAAAAATACA 133

RESULT 12
BF819673/c
LOCUS
DEFINITION
MR1-RT0028-101100-002-e10 RT0028 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BF819673
VERSION
BF819673.1 GI:12157676
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 246)
AUTHORS
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
20202663
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR1&t2=MR1-RT0028-
101100-002-e10&t3=2000-11-10&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 48
High quality sequence stop: 246.
FEATURES
Location/Qualifiers
1..246
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="RT0028"
/dev_stage="Adult"
/note="Organ: Kidney_tumor; Vector: puc18; Site_1: Smar;
Site_2: Smar; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      39 a   61 c   69 g   77 t
ORIGIN
Query Match      83.5%;   Score 42.6;   DB 12;   Length 246;
Best Local Similarity 88.2%;   Pred. No. 0.00091;
Matches 45;   Conservative 1;   Mismatches 5;   Indels 0;   Gaps 0;

```

```

Qy 1 TCAAGACACCGCTGCAGAACTTGGAGAACCGGTCTCTCTACAAAAATACA 51
|||||
Db 219 TCAAGACACCGCTGGGCAACATGGCAAAACCCCTGCTCTCTACAAAAATACA 169

RESULT 13
AI970561
LOCUS
DEFINITION
wrl2a09.xl1 NCI_CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2481304 3'
similar to contains Alu repetitive element; contains LTR5 repetitive
element ;, mRNA sequence.
ACCESSION
AI970561
VERSION
AI970561.1 GI:5767387
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 268)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbrrp/image/image.html
Insert length: 546 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 267.
FEATURES
Location/Qualifiers
1..268
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2481304"
/clone_lib="NCI_CGAP_Lu19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT      93 a   54 c   66 g   55 t
ORIGIN
Query Match      83.5%;   Score 42.6;   DB 9;   Length 268;
Best Local Similarity 88.2%;   Pred. No. 0.00092;
Matches 45;   Conservative 1;   Mismatches 5;   Indels 0;   Gaps 0;

Qy 1 TCAAGACACCGCTGCAGAACTTGGAGAACCGGTCTCTACAAAAATACA 51
|||||
Db 207 TCAAGACACCGCTGGGCAACATGGCAAAACCCCTGCTCTCTACAAAAATACA 257

RESULT 14
BG990678
LOCUS
DEFINITION
RC4-HT1092-250101-022-b05 HT1092 Homo sapiens cDNA, mRNA sequence.

```



ACCESSION BG990678  
VERSION BG990678.1  
KEYWORDS GI:114394748  
SOURCE EST.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 274)  
REFERENCE Dias Neto,E., Garcia Correa,R., Verjovsky-Almeida,S., Briones,M.R.,  
AUTHORS Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC4&st2=RC4-HT1092-  
250101-022-b05&st3=2001-01-25&st4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 15  
High quality sequence stop: 196.  
Location/Qualifiers  
1..274  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="HT1092"  
/dev\_stage="Adult"  
/note="Organ: head\_neck; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 70 a 75 c 81 g 48 t  
ORIGIN  
Query Match 83.5%; Score 42.6; DB 13; Length 274;  
Best Local Similarity 88.2%; Pred. No. 0.00092;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TCAAGACCAGCTGGACAACCTTGGAGAACCSSGGTCTCTACAAAAATACA 51  
|||||  
Db 75 TCAAGACCAGCTGGACAACATGGCAAAACCCGCTCTCTACTAAAAATACA 125  
|||||  
RESULT 15  
AW103010/c 303 bp mRNA linear EST 19-OCT-1999  
LOCUS xd4hl1.x1 NCI\_CGAP\_Ov23 Homo sapiens cDNA clone IMAGE:2596389 3'  
DEFINITION similar to contains Alu repetitive element;; mRNA sequence.  
ACCESSION AW103010  
VERSION AW103010.1  
KEYWORDS GI:6073625  
SOURCE EST.  
ORGANISM human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 303)  
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
TITLE

JOURNAL Tumor Gene Index  
COMMENT Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Seq primer: -400P from Gibco  
High quality sequence stop: 297.  
Location/Qualifiers  
1..303  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:2596389"  
/clone\_lib="NCI\_CGAP\_Ov23"  
/tissue\_type="tumor, 5 pooled (see description)"  
/lab\_host="DH10B"  
/note="Organ: ovary; Vector: pCMV-SPORT6; Site\_1: SalI;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.35 kb. Tumor types include: mixed  
Mullerian tumor, papillary serous, clear cell, spindle  
cell. All are primary tumors, metastasis positive. Life  
Technologies catalog #: 11534-013"  
BASE COUNT 55 a 79 c 69 g 100 t  
ORIGIN  
Query Match 83.5%; Score 42.6; DB 10; Length 303;  
Best Local Similarity 88.2%; Pred. No. 0.00092;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TCAAGACCAGCTGGACAACCTTGGAGAACCSSGGTCTCTACAAAAATACA 51  
|||||  
Db 177 TCAAGACCAGCTGGCAACATGGAGAACCCTCTCTCTACAAAAATACA 127  
|||||  
Search completed: February 11, 2003, 03:31:06  
Job time : 151.133 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 21:12:29 ; Search time 3.87546 Seconds  
(without alignments)  
4035.781 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_5\_55

Perfect score: 51  
Sequence: 1 tcaagaccagcctggacaac.....sggtctctacaaaaataca 51

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	42.6	83.5	12047	2	US-09-022-461-1
C 2	42.6	83.5	12047	4	US-09-033-556-3
C 3	41	80.4	16389	4	US-09-741-154-3
C 4	41	80.4	17327	1	US-07-906-871-15
C 5	41	80.4	36159	4	US-09-749-588-3
C 6	41	80.4	45546	4	US-09-146-053-6
C 7	41	80.4	59065	4	US-09-813-817-3
C 8	41	80.4	59065	4	US-09-978-197-3
C 9	41	80.4	162450	4	US-09-345-882-1
C 10	39.4	77.3	282	1	US-08-133-629-8
C 11	39.4	77.3	327	3	US-08-741-406-6
C 12	39.4	77.3	327	3	US-09-024-472-6
C 13	39.4	77.3	361	4	US-09-018-584A-16
C 14	39.4	77.3	372	4	US-09-018-584A-15
C 15	39.4	77.3	488	4	US-09-385-982-471
C 16	39.4	77.3	901	1	US-08-832-883-65
C 17	39.4	77.3	901	2	US-08-832-883-65
C 18	39.4	77.3	1278	2	US-08-909-965C-4
C 19	39.4	77.3	1659	4	US-08-927-165A-5
C 20	39.4	77.3	2612	4	US-09-484-970B-142
C 21	39.4	77.3	2896	2	US-08-709-923-1
C 22	39.4	77.3	3565	1	US-08-578-649-3
C 23	39.4	77.3	3627	4	US-09-323-873A-6
C 24	39.4	77.3	3865	1	US-08-832-883-48
C 25	39.4	77.3	3865	2	US-08-832-877-48
C 26	39.4	77.3	11358	5	PCR-US93-06251-23
C 27	39.4	77.3	12394	4	US-09-488-856A-10

28	39.4	77.3	14796	4	US-08-975-080-35	Sequence 35, Appl
29	39.4	77.3	14796	4	US-09-630-706-10	Sequence 10, Appl
30	39.4	77.3	14796	4	US-09-496-694B-3	Sequence 3, Appl
C 31	39.4	77.3	18073	4	US-09-078-294-12	Sequence 12, Appl
32	39.4	77.3	28720	4	US-09-341-587-7	Sequence 7, Appl
33	39.4	77.3	35060	3	US-08-814-095-7	Sequence 7, Appl
C 34	39.4	77.3	35060	3	US-08-814-095-7	Sequence 7, Appl
C 35	39.4	77.3	40000	4	US-09-780-049-18	Sequence 18, Appl
C 36	39.4	77.3	43950	4	US-09-735-934A-3	Sequence 3, Appl
C 37	39.4	77.3	50000	4	US-09-146-053-3	Sequence 3, Appl
C 38	39.4	77.3	62804	4	US-09-800-960-3	Sequence 3, Appl
C 39	39.4	77.3	70000	4	US-09-851-896-3	Sequence 3, Appl
C 40	39.4	77.3	70000	4	US-09-851-896-3	Sequence 3, Appl
C 41	39.4	77.3	84495	4	US-09-797-906-3	Sequence 3, Appl
C 42	39.4	77.3	87350	3	US-08-781-891-79	Sequence 79, Appl
C 43	39.4	77.3	87543	4	US-09-791-211-3	Sequence 3, Appl
C 44	39.4	77.3	98844	4	US-09-791-211-10	Sequence 10, Appl
C 45	39.4	77.3	99500	4	US-09-798-096-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-09-022-461-1/c  
; Sequence 1, Application US/09022461  
; Patent No. 5964371  
; GENERAL INFORMATION:  
; APPLICANT: HENDERSON, Daniel R.  
; APPLICANT: SCHUUR, Eric R.  
; APPLICANT: LAMPARSKI, Henry G.  
; APPLICANT: YU, De Chao  
; TITLE OF INVENTION: PROSTATE CANCER DRUG SCRE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/022,461  
; FILING DATE: 12-FEB-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/906,192  
; FILING DATE: 04-AUG-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Catherine, Polizzi M  
; REGISTRATION NUMBER: 40,130  
; REFERENCE/DOCKET NUMBER: 34802-20003.21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-813-5600  
; TELEFAX: 415-494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12047 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-09-022-461-1

Query Match 83.5%; Score 42.6; DB 2; Length 12047;  
Best Local Similarity 88.2%; Pred. No. 3,le-07;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACAGCCTGGACAACCTTGGAGAACCGGTCTCTCTACAAAAAATACA 51  
|||||  
Db 1985 TCAAGACAGCCTGGCCAACATGGCAAAACCCCGTCTCTACAAAAAATACA 1935

RESULT 2  
US-09-033-556-3/c  
; Sequence 3, Application US/09033556  
; Patent No. 6432700  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Daniel R.  
; APPLICANT: Yu, De Chao  
; TITLE OF INVENTION: ADENOVIRUS VECTORS CONTAINING  
; TITLE OF INVENTION: HETEROLOGOUS TRANSCRIPTION REGULATORY ELEMENTS AND METHODS  
; TITLE OF INVENTION: OF USING SAME  
; NUMBER OF SEQUENCES: 41  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/033,556  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Catherine, Polizzi M  
; REGISTRATION NUMBER: 40,130  
; REFERENCE/DOCKET NUMBER: 34802-20010.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-813-5600  
; TELEFAX: 650-494-0792  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12047 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-09-033-556-3

Query Match 83.5%; Score 42.6; DB 4; Length 12047;  
Best Local Similarity 88.2%; Pred. No. 3.1e-07;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 TCAAGACAGCCTGGACAACCTTGGAGAACCGGTCTCTCTACAAAAAATACA 51  
|||||  
Db 1985 TCAAGACAGCCTGGCCAACATGGCAAAACCCCGTCTCTACAAAAAATACA 1935

RESULT 3  
US-09-741-154-3/c  
; Sequence 3, Application US/09741154  
; Patent No. 6437110  
; GENERAL INFORMATION:  
; APPLICANT: BEASLEY, Ellen M. et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001061  
; CURRENT APPLICATION NUMBER: US/09/741,154  
; CURRENT FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 16389  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-741-154-3  
Query Match 80.4%; Score 41; DB 4; Length 16389;  
Best Local Similarity 86.3%; Pred. No. 1.4e-06;  
Matches 44; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 1 TCAAGACAGCCTGGACAACCTTGGAGAACCGGTCTCTCTACAAAAAATACA 51  
|||||  
Db 455 TCAAGACAGCCTGGCCAACATGGCAAAACCCCGTCTCTCTACAAAAAATACA 405  
RESULT 4  
US-07-906-871-15/c  
; Sequence 15, Application US/07906871  
; Patent No. 5340739  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; APPLICANT: Avraham, Shalom  
; TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC  
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY ELEMENTS OF SERGLYCIN AND USES  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1225 Connecticut Avenue, N.W., Suite 300  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/906,871  
; FILING DATE: 19920103  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/816,289  
; FILING DATE: 03 JAN 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/635,544  
; FILING DATE: 18-JAN-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US89/03051  
; FILING DATE: 13-JUL-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/224,035  
; FILING DATE: 13-JUL-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cimbala, Michele A  
; REGISTRATION NUMBER: 33,851  
; REFERENCE/DOCKET NUMBER: 0627.2830004  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)833-7533  
; TELEFAX: (202)833-8716  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 17327 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: both  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 621..753

```
;
; FEATURE:
; NAME/KEY: intron
; LOCATION: 754..9596
; FEATURE:
; NAME/KEY: exon
; LOCATION: 9597..9744
; FEATURE:
; NAME/KEY: intron
; LOCATION: 9745..16396
; FEATURE:
; NAME/KEY: exon
; LOCATION: 16397..17327
;
US-07-906-871-15

Query Match      80.4%; Score 41; DB 1; Length 17327;
Best Local Similarity 86.3%; Pred. No. 1.4e-06;
Matches 44; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACCTTGGAGAACCSCGGTCTCTACAAAAAATACA 51
|||||
Db 13706 TCAAGACCAGCCTGGTCAACATGGCAAAACCCGCTCTCTACCAAAAAATACA 13656

RESULT 5
US-09-749-588-3
; Sequence 3, Application US/09749588
; Patent No. 6423521
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001068
; CURRENT APPLICATION NUMBER: US/09/749,588
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36159
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(36159)
; OTHER INFORMATION: n = A,T,C or G
US-09-749-588-3

Query Match      80.4%; Score 41; DB 4; Length 36159;
Best Local Similarity 86.3%; Pred. No. 1.7e-06;
Matches 44; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACCTTGGAGAACCSCGGTCTCTACAAAAAATACA 51
|||||
Db 32026 TCAAGACCAGCCTGGCCAAACATGGCACACCCTGTCTCTACTAAAAAATACA 32076

RESULT 6
US-09-146-053-6
; Sequence 6, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Sprinkle, Terry Joe Curtis
; APPLICANT: Venema, Richard C.
; TITLE OF INVENTION: Human Aminopeptidase P Gene
; FILE REFERENCE: MCG103
; CURRENT APPLICATION NUMBER: US/09/146,053A
; CURRENT FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/057,854
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
```

```
;
; LENGTH: 45546
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-053-6

Query Match      80.4%; Score 41; DB 4; Length 45546;
Best Local Similarity 86.3%; Pred. No. 1.8e-06;
Matches 44; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACCTTGGAGAACCSCGGTCTCTACAAAAAATACA 51
|||||
Db 39107 TCAAGACCAGCCTGGCCAAACATGTGAACCCCTGTCTCTACTAAAAAATACA 39157

RESULT 7
US-09-813-817-3/C
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-813-817-3

Query Match      80.4%; Score 41; DB 4; Length 59065;
Best Local Similarity 86.3%; Pred. No. 1.9e-06;
Matches 44; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACCTTGGAGAACCSCGGTCTCTACAAAAAATACA 51
|||||
Db 49921 TCAAGACCAGCCTGGCCAAACATGGCAAAACCCCTGTCTCTACTAAAAAATACA 49871

RESULT 8
US-09-978-197-3/C
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001178DIV
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-978-197-3

Query Match      80.4%; Score 41; DB 4; Length 59065;
Best Local Similarity 86.3%; Pred. No. 1.9e-06;
Matches 44; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACCTTGGAGAACCSCGGTCTCTACAAAAAATACA 51
|||||
Db 49921 TCAAGACCAGCCTGGCCAAACATGGCAAAACCCCTGTCTCTACTAAAAAATACA 49871
```

```
RESULT 9
US-09-345-882-1
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; FILE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93714
; OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97122
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97152
; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99098
; OTHER INFORMATION: 5-130-257 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99117
; OTHER INFORMATION: 5-130-276 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103806
; OTHER INFORMATION: 5-131-395 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106940
; OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108106
; OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108149
; OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108308
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108471
; OTHER INFORMATION: 5-136-174 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134134
; OTHER INFORMATION: 5-140-120 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134362
; OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146328
; OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146345
; OTHER INFORMATION: 5-143-84 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 150329
; OTHER INFORMATION: 5-145-24 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 160031
; OTHER INFORMATION: 5-148-352 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90819..90865
; OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93690..93736
; OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97099..97145
; OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97099..97145
; OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
```











GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 03:31:30 ; Search time 4.62902 Seconds  
(without alignments)  
5222.300 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_5\_55

Perfect score: 51  
Sequence: 1 tcaagaccagcctggacac.....sggtctacaaaaataca 51

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	42.6	83.5	401	9	US-09-946-807-952
C 2	42.6	83.5	401	10	US-09-795-668-952
C 3	42.6	83.5	401	10	US-09-795-686-952
C 4	42.6	83.5	4962	10	US-09-764-869-1902
C 5	42.6	83.5	4962	10	US-09-764-869-1903
C 6	42.6	83.5	12047	10	US-09-392-822-5
C 7	42.6	83.5	12047	10	US-09-873-228-1
C 8	42.6	83.5	56737	10	US-09-782-378A-17
C 9	42.6	83.5	116592	10	US-09-818-512-3
C 10	42.6	83.5	174493	9	US-10-238-709-3
C 11	42.6	83.5	174493	10	US-09-804-71A-3
C 12	41.4	81.2	1717	10	US-09-764-869-1636
C 13	41	80.4	242	10	US-09-867-701-8259
C 14	41	80.4	317	10	US-09-908-711-22
C 15	41	80.4	419	10	US-09-867-701-8232
C 16	41	80.4	464	10	US-09-867-701-10666
C 17	41	80.4	475	10	US-09-867-701-6866
C 18	41	80.4	476	10	US-09-867-701-10182
C 19	41	80.4	497	10	US-09-867-701-2477

C 20	41	80.4	552	10	US-09-867-701-2467	Sequence 2467, Ap
C 21	41	80.4	1762	10	US-09-893-348-4	Sequence 4, Appli
C 22	41	80.4	2446	10	US-09-822-849A-418	Sequence 418, App
C 23	41	80.4	4978	10	US-09-764-887-355	Sequence 355, App
C 24	41	80.4	6371	10	US-09-876-889-345	Sequence 345, App
C 25	41	80.4	6640	10	US-09-764-847-1077	Sequence 1077, Ap
C 26	41	80.4	8369	10	US-09-764-903-67	Sequence 67, Appl
C 27	41	80.4	8974	10	US-09-764-847-1078	Sequence 1078, Ap
C 28	41	80.4	9375	10	US-09-764-877-3449	Sequence 3449, Ap
C 29	41	80.4	9558	10	US-09-764-877-3645	Sequence 3645, Ap
C 30	41	80.4	10923	10	US-09-764-847-1079	Sequence 1079, Ap
C 31	41	80.4	11869	10	US-09-764-869-2292	Sequence 2292, Ap
C 32	41	80.4	13176	10	US-09-764-869-1657	Sequence 1657, Ap
C 33	41	80.4	13996	10	US-09-764-870-602	Sequence 602, App
C 34	41	80.4	14001	10	US-09-764-870-601	Sequence 601, App
C 35	41	80.4	14417	9	US-09-860-670-251	Sequence 251, App
C 36	41	80.4	14426	9	US-09-860-670-249	Sequence 249, App
C 37	41	80.4	14426	9	US-09-860-670-252	Sequence 252, App
C 38	41	80.4	14448	9	US-09-860-670-250	Sequence 250, App
C 39	41	80.4	14451	9	US-09-860-670-253	Sequence 253, App
C 40	41	80.4	15044	10	US-09-764-869-1790	Sequence 1790, Ap
C 41	41	80.4	15046	10	US-09-764-869-1791	Sequence 1791, Ap
C 42	41	80.4	15849	10	US-09-880-107-2362	Sequence 2362, Ap
C 43	41	80.4	17335	10	US-09-764-847-1280	Sequence 1280, Ap
C 44	41	80.4	19882	10	US-09-764-847-1281	Sequence 1281, Ap
C 45	41	80.4	23580	10	US-09-764-860-990	Sequence 990, App

ALIGNMENTS

RESULT 1

US-09-946-807-952/c  
; Sequence 952, Application US/09946807  
; Patent No. US20020165144A1  
; GENERAL INFORMATION:  
; APPLICANT: Stefansson, Hreinn  
; APPLICANT: Steinthorsdottir, Valgerdur  
; APPLICANT: Gulcher, Jeffrey R.  
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE  
; FILE REFERENCE: 2345, 2004-001  
; CURRENT APPLICATION NUMBER: US/09/946, 807  
; CURRENT FILING DATE: 2001-09-05  
; PRIOR APPLICATION NUMBER: US/09/795, 668  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 09/515,716  
; PRIOR FILING DATE: 2000-02-28  
; NUMBER OF SEQ ID NOS: 1531  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 952  
; LENGTH: 401  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-946-807-952

Query Match 83.5%; Score 42.6; DB 9; Length 401;  
Best Local Similarity 88.2%; Pred. No. 3.2e-07;  
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCACCTGGACAACTTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 84 TCAAGACCACCTGGGCAACATGAAAAACCTGTCTCTACAAAAATACA 34  
|||||

RESULT 2

US-09-795-668-952/c  
; Sequence 952, Application US/09795668  
; Patent No. US20020045577A1  
; GENERAL INFORMATION:  
; APPLICANT: Stefansson, Hreinn  
; APPLICANT: Steinthorsdottir, Valgerdur  
; APPLICANT: Gulcher, Jeffrey R.  
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE

```
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/795,668
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 952
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-668-952

Query Match      83.5%; Score 42.6; DB 10; Length 401;
Best Local Similarity 88.2%; Pred. No. 3.2e-07;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAAGCCTGGACAACCTTGGAGAAGACCCGGTCTCTACAAAAAATACA 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 TCAAGACCAAGCCTGGCAACATGGAAAAACCCCTGTCTCTACCAAAAAATACA 34

RESULT 3
US-09-795-686-952/c
; Sequence 952, Application US/09795686
; Patent No. US20020094954A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2005-001
; CURRENT APPLICATION NUMBER: US/09/795,686
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,715
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 952
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-686-952

Query Match      83.5%; Score 42.6; DB 10; Length 401;
Best Local Similarity 88.2%; Pred. No. 3.2e-07;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAAGCCTGGACAACCTTGGAGAAGACCCGGTCTCTACAAAAAATACA 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 84 TCAAGACCAAGCCTGGCAACATGGAAAAACCCCTGTCTCTACCAAAAAATACA 34

RESULT 4
US-09-764-869-1902
; Sequence 1902, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1902
; LENGTH: 4962
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1902

Query Match      83.5%; Score 42.6; DB 10; Length 4962;
Best Local Similarity 88.2%; Pred. No. 10;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAAGCCTGGACAACCTTGGAGAAGACCCGGTCTCTACAAAAAATACA 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1985 TCAAGACCAAGCCTGGCAACATGGCAAAACCCCGTCTCTACAAAAAATACA 1935

RESULT 5
US-09-764-869-1903
; Sequence 1903, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1903
; LENGTH: 4962
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1903

Query Match      83.5%; Score 42.6; DB 10; Length 4962;
Best Local Similarity 88.2%; Pred. No. 6.4e-07;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAAGCCTGGACAACCTTGGAGAAGACCCGGTCTCTACAAAAAATACA 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2690 TCAAGACCAAGCCTGGCAACATGGTGAACCCCTGTCTCTACAAAAAATACA 2740

RESULT 6
US-09-392-822-5/c
; Sequence 5, Application US/09392822
; Patent No. US20010053352A1
; GENERAL INFORMATION:
; APPLICANT: Yu, De Chao
; APPLICANT: Henderson, Daniel
; TITLE OF INVENTION: Adenovirus Vectors Containing Cell
; TITLE OF INVENTION: Status-Specific Response Elements and Methods of Use Thereof
; FILE REFERENCE: 348022001200
; CURRENT APPLICATION NUMBER: US/09/392,822
; CURRENT FILING DATE: 1999-09-09
; EARLIER APPLICATION NUMBER: 60/099,791
; EARLIER FILING DATE: 1998-09-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 12047
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-392-822-5

Query Match      83.5%; Score 42.6; DB 10; Length 12047;
Best Local Similarity 88.2%; Pred. No. 8.1e-07;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAAGCCTGGACAACCTTGGAGAAGACCCGGTCTCTACAAAAAATACA 51
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1985 TCAAGACCAAGCCTGGCAACATGGCAAAACCCCGTCTCTACAAAAAATACA 1935

RESULT 7
US-09-875-228-1/c
; Sequence 1, Application US/09875228
; Patent No. US20020136707A1
; GENERAL INFORMATION:
; APPLICANT: Yu, D.
```

```
; APPLICANT: Henderson, D.R.
; APPLICANT: Schuur, E.R.
; TITLE OF INVENTION: A HUMAN GLANDULAR KALLIKREIN ENHANCER, VECTORS COMPRISING THE
; TITLE OF INVENTION: ENHANCER AND METHODS OF USE THEREOF
; FILE REFERENCE: 348022000900
; CURRENT APPLICATION NUMBER: US/09/875,228
; CURRENT FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 09/127,834
; PRIOR FILING DATE: 1998-08-03
; PRIOR APPLICATION NUMBER: 60/076,545
; PRIOR FILING DATE: 1998-03-02
; PRIOR APPLICATION NUMBER: 60/054,523
; PRIOR FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 12047
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-875-228-1

Query Match      83.5%; Score 42.6; DB 10; Length 12047;
Best Local Similarity 88.2%; Pred. No. 8.1e-07;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCSGGTCTCTACAAAAAATACA 51
|||||
Db 1985 TCAAGACCAGCCTGGCAACATGGCAAAACCCGCTCTCTACAAAAAATACA 1935

RESULT 8
US-09-782-378A-17
; Sequence 17, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 60/237,747
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 56737
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-782-378A-17

Query Match      83.5%; Score 42.6; DB 10; Length 56737;
Best Local Similarity 88.2%; Pred. No. 1.2e-06;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCSGGTCTCTACAAAAAATACA 51
|||||
Db 56597 TCAAGACCAGCCTGGCAACATGGCAGAACCCCGCTCTCTACTAAAAAATACA 56647

RESULT 9
US-09-818-512-3
; Sequence 3, Application US/09818512
; Patent No. US20020142416A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001192
; CURRENT APPLICATION NUMBER: US/09/818,512
```

```
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(116592)
; OTHER INFORMATION: n = A,T,C or G
US-09-818-512-3

Query Match      83.5%; Score 42.6; DB 10; Length 116592;
Best Local Similarity 88.2%; Pred. No. 1.5e-06;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCSGGTCTCTACAAAAAATACA 51
|||||
Db 56955 TCAAGACCAGCCTGGACAACTTGGCAAAACCCCATCTCTACAAAAAATACA 57005

RESULT 10
US-10-238-709-3/C
; Sequence 3, Application US/10238709
; Publication No. US20030022340A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marlon et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164DIV
; CURRENT APPLICATION NUMBER: US/10/238,709
; CURRENT FILING DATE: 2002-09-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(174493)
; OTHER INFORMATION: n = A,T,C or G
US-10-238-709-3

Query Match      83.5%; Score 42.6; DB 9; Length 174493;
Best Local Similarity 88.2%; Pred. No. 1.7e-06;
Matches 45; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 TCAAGACCAGCCTGGACAACTTGGAGAACCCSGGTCTCTACAAAAAATACA 51
|||||
Db 67754 TCAAGACCAGCCTGGCAACATGGCAAAACCCGTCTCTACAAAAAATACA 67704

RESULT 11
US-09-804-471A-3/C
; Sequence 3, Application US/09804471A
; Patent No. US20020132322A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marlon et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
```

[illegible]

; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US01/01312  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/209,467  
; PRIOR FILING DATE: 2000-06-07  
; NUMBER OF SEQ ID NOS: 167  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 317  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (138)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (160)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-908-711-22

Query Match 80.4%; Score 41; DB 10; Length 317;  
Best Local Similarity 86.3%; Pred. No. 1.2e-06;  
Matches 44; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGACAACCTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 75 TCAAGACCGCTGGCCAACATGGCAAAACCCGTCTCTACTAAATACA 25  
|||||

RESULT 15

US-09-867-701-8232/c  
; Sequence 8232, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8232  
; LENGTH: 419  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-867-701-8232

Query Match 80.4%; Score 41; DB 10; Length 419;  
Best Local Similarity 86.3%; Pred. No. 1.3e-06;  
Matches 44; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCAAGACCGCTGGACAACCTGGAGAACCCGGTCTCTACAAAAATACA 51  
|||||  
Db 264 TCAAGACCGCTGGCCAACATGGCAAAACCCGTCTCTACTAAATACA 214  
|||||

Search completed: February 11, 2003, 09:07:51  
Job time : 61.629 secs





GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:43:59 ; Search time 16.6596 seconds  
(without alignments)  
5542.256 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_175\_215

Perfect score: 41  
Sequence: 1 cctatctctactgaataatay.....aaaagctagacgtggtggca 41

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1135999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.6	99.0	1669	22	AAH26169 Human cytochrome P
2	40.6	99.0	1669	22	AAH26179 Human cytochrome P
3	40.6	99.0	1680	24	AAD34214 Human CYP2D6 gene
4	40.6	99.0	9432	24	AAD34213 Human cytochrome P
5	32.6	79.5	58708	22	AAK64739 Human immune/haema
6	31.6	77.1	9704	24	ABL60777 Genomic DNA of an
7	31	75.6	2295	22	AA105137 Human reproductive
8	31	75.6	31730	22	AA137445 Human musculoskele
9	31	75.6	86080	24	ABQ88164 Human osteoblast d

10	31	75.6	86080	24	ABK83561 Human cDNA differe
c 11	31	75.6	167343	24	ABL64403 Stomach cancer rel
c 12	31	75.6	167343	24	ABL67239 Thyroid cancer rel
13	30.6	74.6	264	22	AA105992 Human restructive
14	30.6	74.6	264	23	ABL98556 Human testicular a
c 15	30	73.2	19315	22	AA528883 Human immunoglobul
c 16	30	73.2	103747	24	ABQ88139 Human HBM gene reg
c 17	29.8	72.7	26928	22	ABA82620 Human high bone ma
c 18	29.8	72.7	26928	24	ABK22779 Human SNP oligonuc
c 19	29.4	71.7	50	22	AA128464 Human cancer relat
20	29.4	71.7	627	24	ABN63715 Human musculoskele
c 21	29.4	71.7	10514	22	ABL37105 Colon adenocarcino
c 22	29	70.7	400	24	ABL61718 Colon adenocarcino
c 23	29	70.7	400	24	ABL62279 Human cDNA clone (
c 24	28.8	70.2	559	22	AAH09284 Human prostate exp
c 25	28.8	70.2	579	23	ABV55775 Novel human polynu
c 26	28.6	69.8	606	22	ABF65385 Human prostate exp
27	28.4	69.3	548	23	ABV44644 Human prostate exp
28	28.4	69.3	681	23	ABV05598 Human immune/haema
29	28.4	69.3	6965	22	AAK80192 Human immune/haema
30	28.4	69.3	6976	22	AAK80193 CYB5RP fatty acid
c 31	28.4	69.3	18402	21	AAZ93705 Human immune/haema
c 32	28.4	69.3	39887	22	NAK79153 Human immune/haema
c 33	28.4	69.3	39887	22	AAK81263 Human immune/haema
c 34	28.4	69.3	54297	22	AAK70625 Human immune/haema
c 35	28.2	68.8	274	22	AAK77747 Human immune/haema
c 36	28.2	68.8	274	22	AAK78708 Human breast cance
37	28.2	68.8	416	22	AA125612 Human breast cance
38	28.2	68.8	416	22	AA125613 Human breast cance
39	28.2	68.8	430	22	AA118007 Human breast cance
40	28.2	68.8	430	22	AA118008 Human breast cance
41	28.2	68.8	446	22	AA107867 Human breast cance
42	28.2	68.8	465	22	AA107866 Human breast cance
43	28	68.3	551	23	ABV58508 Human prostate exp
44	28	68.3	1180	21	AAH81674 N. meningitidis pa
c 45	28	68.3	13559	22	ABAI5144 Human nervous syst

ALIGNMENTS

RESULT 1  
AAH26169  
ID AAH26169 standard; DNA; 1669 BP.  
AC AC  
AAH26169;  
XX  
XX  
DT 17-SEP-2001 (first entry)  
XX  
DE Human cytochrome P450 CYP2D6 gene promoter region.  
XX  
KW Cytochrome P450; CYP2D6; promoter; drug metabolism; human;  
KW diagnosis; therapy; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT primer\_bind Location/Qualifiers  
FT complement (14...36)  
FT /\*tag= a  
FT /note= "amplification primer upf14"  
FT primer\_bind 337...358  
FT /\*tag= b  
FT /note= "sequence primer R1"  
FT primer\_bind 493...514  
FT /\*tag= c  
FT /note= "sequencing primer R2"  
FT primer\_bind complement (565..577)  
FT /\*tag= d  
FT /note= "sequencing primer F2"  
FT primer\_bind 602..620  
FT /\*tag= e  
FT /note= "sequencing primer R3"  
FT primer\_bind complement (968..988)

FT /\*tag= f  
 FT /note= "sequencing primer F3"  
 FT 1124..1143  
 FT /\*tag= g  
 FT /note= "sequencing primer R4"  
 FT 1605..1623  
 FT /\*tag= h  
 FT /note= "sequencing primer R5"  
 FT 1650..1669  
 FT /\*tag= i  
 FT /note= "amplification primer upr1669"  
 FT 1532..1619  
 FT /\*tag= j  
 FT 1620..1669  
 FT /\*tag= k  
 FT /partial  
 FT /note= "5' region of CYP2D6 coding region"

XX WO200155432-A2.

XX 02-AUG-2001.

XX 30-JAN-2001; 2001WO-EP00954.

XX 31-JAN-2000; 2000EP-0101889.

XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX Raimundo S, Zanger U;

XX WPI; 2001-457734/49.

XX A polynucleotide capable of hybridizing to CYP2D6 promoter useful for  
 PT the optimization of drug therapies using substrates of cytochrome P-450  
 PT -

XX Claim 1; Fig 1; 41pp; English.

XX The present sequence is that of the promoter region of the human  
 CC cytochrome P450 CYP2D6 gene. The promoter region was amplified  
 CC by PCR from leucocyte DNA of over 50 individuals, and sequenced.  
 CC 8 Previously unknown single nucleotide polymorphisms (SNP) were  
 CC identified. These were at: base 36 (base -1584 according to the  
 CC human Cytochrome P450 Allele Nomenclature), where the SNP was C to  
 CC G, occurring at an estimated frequency of approximately 20% in the  
 CC whole population, and resulting in increased enzyme activity;  
 CC position 194 (-1426), C to T, approximately 20% frequency; neutral  
 CC function; position 385 (-1235), A to G, approximately 50% frequency,  
 CC neutral function; position 620 (-1000), G to A, approximately 20%  
 CC frequency; neutral function; position 880 (-740), C to T,  
 CC approximately 30% frequency; unknown function; position 940 (-680),  
 CC G to A, approximately 30% frequency; unknown function; 1255 (-365),  
 CC G to A, rare, unknown function; and 1298 (-322), T to C, rare,  
 CC unknown function. The C to G mutation at -1584 bp is strongly  
 CC associated with lower metabolic ratios, and a molecular variant  
 CC Polynucleotide having G at this position is claimed (see AAH26179).  
 CC The invention provides a method of diagnosing a disorder related to  
 CC reduced or enhanced capacity for clearance of CYP2D6 substrates  
 CC (antiarrhythmic, beta-adrenergic receptor-antagonist, tricyclic  
 CC antidepressant, selective serotonin reuptake inhibitor, neuroleptic,  
 CC opiate, cytostatic or amphetamine), or susceptibility to such a  
 CC disorder, by determining the presence of a mutation in the CYP2D6  
 CC promoter. The strong association of the common C to G mutation at  
 CC -1584 bp with increased enzyme activity significantly improves the  
 CC correlation between genotype and phenotype in the CYP2D6 polymorphism.  
 CC Testing for the mutation will allow the identification of intermediate  
 CC metabolizers and therefore allow quantitative predictions to be made  
 CC on in vivo drug metabolism capacity, thus providing a very potent  
 CC tool for improving the therapy of diseases with drugs that are  
 CC targets of the CYP2D6 gene product.

XX Sequence 1669 BP; 413 A; 376 C; 534 G; 338 T; 8 Other;

Query Match 99.0%; Score 40.6; DB 22; Length 1669;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 41  
 |||||  
 Db 175 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 215

RESULT 2

AAH26179

ID AAH26179 standard; DNA; 1669 BP.

XX AAH26179;

XX 17-SEP-2001 (first entry)

XX Human cytochrome P450 CYP2D6 gene promoter (G mutation at -1584 bp).

XX Cytochrome P450; CYP2D6; promoter; drug metabolism; human;  
 KW diagnosis; therapy; single nucleotide polymorphism; ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
 FT variation replace(36.G)  
 FT /\*tag= a

FT /frequency= "20%"

FT 1532..1619

FT /\*tag= b

FT 1620..1669

FT /\*tag= c

FT /partial

FT /note= "5' region of CYP2D6 coding region"

XX WO200155432-A2.

XX 02-AUG-2001.

XX 30-JAN-2001; 2001WO-EP00954.

XX 31-JAN-2000; 2000EP-0101889.

XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX Raimundo S, Zanger U;

XX WPI; 2001-457734/49.

XX A polynucleotide capable of hybridizing to CYP2D6 promoter useful for  
 PT the optimization of drug therapies using substrates of cytochrome P-450  
 PT -

XX Claim 1(a); Page -; 41pp; English.

XX The present sequence is that of the promoter region of the human  
 CC cytochrome P450 CYP2D6 gene, which includes G at position 36  
 CC (base -1584 according to the Human Cytochrome P450 Allele  
 CC Nomenclature). The presence of C at position -1584 bp is a marker  
 CC for low enzyme activity, whereas there is strong association of G  
 CC at position -1584 bp with increased enzyme activity. The C to G  
 CC single nucleotide polymorphism occurs in approximately 20% of the  
 CC population. The invention provides a method of diagnosing a  
 CC disorder related to reduced or enhanced capacity for clearance of  
 CC CYP2D6 substrates (antiarrhythmic, beta-adrenergic receptor  
 CC antagonist, tricyclic antidepressant, selective serotonin reuptake  
 CC inhibitor, neuroleptic, opiate, cytostatic or amphetamine), or  
 CC susceptibility to such a disorder, by determining the presence of  
 CC a mutation in the CYP2D6 promoter. The novel variant forms of the  
 CC CYP2D6 gene provided by the invention provide the potential for the  
 CC development of a pharmacodynamic profile of drugs for a given  
 CC patient. The finding and characterization of variations in the  
 CC CYP2D6 gene, and diagnostic tests for the discrimination of

CC different alleles in human individuals, provide a very potent tool  
CC for improving the therapy of diseases with drugs that are targets  
CC of the CYP2D6 gene production, and whose metabolism is therefore  
CC dependent on CYP2D6 activity.  
CC Note: The present sequence is not shown in the specification but is  
CC derived from the CYP2D6 promoter sequence given in the Sequence  
CC Listing (see AH26169).  
XX  
SQ Sequence 1669 BP; 413 A; 376 C; 535 G; 338 T; 7 other;  
Query Match 99.0%; Score 40.6; DB 22; Length 1669;  
Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGTGGCA 41  
|||||  
Db 175 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGTGGCA 215  
RESULT 3  
AAD34214  
ID AAD34214 standard; DNA; 1680 BP.  
XX  
AC AAD34214;  
XX  
DT 16-JUL-2002 (first entry)  
XX  
DE Human CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX  
KW Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; polymorphism; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 36  
FT /\*tag= a  
FT /note= "Polymorphic site"  
FT 194  
FT /\*tag= b  
FT /note= "Polymorphic site"  
FT 385  
FT /\*tag= c  
FT /note= "Polymorphic site"  
FT 620  
FT /\*tag= d  
FT /note= "Polymorphic site"  
FT 880  
FT /\*tag= e  
FT /note= "Polymorphic site"  
FT 942  
FT /\*tag= f  
FT /note= "Polymorphic site"  
FT 1255  
FT /\*tag= g  
FT /note= "Polymorphic site"  
XX  
WO200218638-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 27-AUG-2001; 2001WO-IB01544.  
XX  
PR 30-AUG-2000; 2000GB-0021286.  
XX  
PA (GEMI-) GEMINI GENOMICS PLC.  
XX  
PI Risinger C, Andersson MK, Lewander T, Olliasson E;  
XX WPI; 2002-329785/36.  
XX  
PT New sequence determination oligonucleotides, useful for detecting

PT polymorphic sites in a 5' flanking region of a CYP2D6 gene, as  
PT hybridization probes, as components of diagnostic assays, or in  
PT ligase-based sequence determination -  
XX  
XX Claim 1; Fig 2; 63pp; English.  
XX  
CC The invention relates to sequence determination oligonucleotides for  
CC detecting polymorphic sites in a 5' flanking region of cytochrome P450  
CC 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many  
CC different xenobiotics. Human CYP2D6 gene is located on chromosome 22.  
CC The oligonucleotides may be used as in situ hybridisation probes, in  
CC ligase-based sequenced determination, as components of diagnostic assays,  
CC as probes in sequence determination methods based on mismatches, as  
CC hybridisation-based diagnostic assays, and as components of diagnostic  
CC microarray. CYP2D6 is useful to predict variations in an individual's  
CC ability to metabolise certain drugs. The present sequence is human  
CC CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX  
SQ Sequence 1680 BP; 413 A; 379 C; 539 G; 342 T; 7 other;  
Query Match 99.0%; Score 40.6; DB 24; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 3.2e-06;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGTGGCA 41  
|||||  
Db 175 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGTGGCA 215  
RESULT 4  
AAD34213  
ID AAD34213 standard; DNA; 9432 BP.  
XX  
AC AAD34213;  
XX  
DT 16-JUL-2002 (first entry)  
XX  
DE Human cytochrome P450 2D6 (CYP2D6) gene.  
XX  
KW Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200218638-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 27-AUG-2001; 2001WO-IB01544.  
XX  
PR 30-AUG-2000; 2000GB-0021286.  
XX  
PA (GEMI-) GEMINI GENOMICS PLC.  
XX  
PI Risinger C, Andersson MK, Lewander T, Olliasson E;  
XX WPI; 2002-329785/36.  
XX  
DR  
XX  
PT New sequence determination oligonucleotides, useful for detecting

CC hybridisation-based diagnostic assays, and as components of diagnostic  
CC microarray. CYP2D6 is useful to predict variations in an individual's  
CC ability to metabolise certain drugs. The present sequence is human  
CC CYP2D6 gene.

XX SQ Sequence 9432 BP; 1964 A; 2647 C; 2976 G; 1845 T; 0 other;  
Query Match 99.0%; Score 40.6; DB 24; Length 9432;  
Best Local Similarity 97.6%; Pred. No. 4.4e-06;  
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATATAAAAAGCTAGACGTGGTGCA 41  
|||||  
Db 175 CCTATCTCTACTGAAATACAAAAAGCTAGACGTGGTGCA 215

RESULT 5  
AAK64739  
ID AAK64739 standard; DNA; 58708 BP.  
XX  
AC AAK64739;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:19551.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.

XX  
XX WO200157182-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01354.

XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216680.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.



```
FT      /*tag= u
FT      /number= "10"
FT      7033..7112
FT      exon
FT      /*tag= v
FT      /number= "11"
FT      7113..7274
FT      intron
FT      /*tag= w
FT      /number= "11"
FT      7275..7335
FT      exon
FT      /*tag= x
FT      /number= "12"
FT      7336..7419
FT      intron
FT      /*tag= y
FT      /number= "12"
FT      7420..7581
FT      exon
FT      /*tag= z
FT      /number= "13"
XX
XX
PN      US6387661-B1.
XX
XX      14-MAY-2002.
XX
XX      23-MAR-2001; 2001US-0814951.
XX
XX      23-MAR-2001; 2001US-0814951.
XX
XX      (PEKE ) PE CORP NY.
XX
XX
PI      Shao W, Yan C, Di Francesco V, Beasley EM;
XX
XX      WPI; 2002-478443/51.
XX      P-PSDB; ABB08102.
XX
XX      Isolated nucleic acid molecules encoding enzymes similar to human
XX      aminocyclase-1, useful as a drug target and diagnostic marker for
XX      cancers e.g. T cell leukemias and ovary, brain or lung cancers -
XX
XX      Disclosure; Fig 3A-E; 43pp; English.
XX
XX      The invention relates to an isolated nucleic acid molecule encoding
XX      enzymes similar to human aminocyclase-1 (ACY-1) (EC 3.5.1.14) (a
XX      metalloprotein cytosolic enzyme). The ACY-1 similar polynucleotide and
XX      encoded peptide sequences can be used as models for the development of
XX      human therapeutic targets, aid in the identification of therapeutic
XX      proteins, and serve as targets for the development of human therapeutic
XX      agents that modulate enzyme activity in cells and tissues that express
XX      the enzyme. ACY-1 has been found to be expressed in humans in the
XX      placenta, T cells from T cell leukemia, ovary, brain, lung and leukocyte,
XX      and therefore may be a drug target for cancer therapy and act as a
XX      diagnostic marker for these cancers. The present sequence represents the
XX      genomic DNA of the enzyme similar to human aminocyclase-1 (ACY-1).
XX
XX      Sequence 9704 BP; 2089 A; 2748 C; 2793 G; 2074 T; 0 other;
SQ
Query Match      77.1%; Score 31.6; DB 24; Length 9704;
Best Local Similarity 85.0%; Pred. No. 0.014;
Matches 34; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY      1 CCTATCTCTACTGAAATATYAAAAGCTAGACGTGGTGCC 40
      || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      8689 CCCGTCCTACTAAAATAACAAAAGCTGGCGTGGTGCC 8650
RESULT 7
ID      AAL05137/c
XX      AAL05137 standard; DNA; 2295 BP.
AC      AAL05137;
XX
XX      21-NOV-2001 (first entry)
XX
XX      Human reproductive system related antigen DNA SEQ ID NO: 7825.
XX
```

```
KW      Human; reproductive system related antigen; reproductive system disorder;
KW      cancer; gene therapy; ds.
XX
XX      Homo sapiens.
OS
PN      WO200155320-A2.
XX
XX      02-AUG-2001.
XX
XX      17-JAN-2001; 2001WO-US01339.
XX
XX      31-JAN-2000; 2000US-0179065.
XX      04-FEB-2000; 2000US-0180628.
XX      24-FEB-2000; 2000US-0184664.
XX      02-MAR-2000; 2000US-0186350.
XX      16-MAR-2000; 2000US-0189874.
XX      17-MAR-2000; 2000US-0190076.
XX      18-APR-2000; 2000US-0198123.
XX      19-MAY-2000; 2000US-0205515.
XX      07-JUN-2000; 2000US-0209467.
XX      28-JUN-2000; 2000US-0214886.
XX      30-JUN-2000; 2000US-0215135.
XX      07-JUL-2000; 2000US-0216647.
XX      07-JUL-2000; 2000US-0216880.
XX      11-JUL-2000; 2000US-0217487.
XX      11-JUL-2000; 2000US-0217496.
XX      14-JUL-2000; 2000US-0218290.
XX      26-JUL-2000; 2000US-0220963.
XX      26-JUL-2000; 2000US-0220964.
XX      14-AUG-2000; 2000US-0224518.
XX      14-AUG-2000; 2000US-0224519.
XX      14-AUG-2000; 2000US-0225213.
XX      14-AUG-2000; 2000US-0225214.
XX      14-AUG-2000; 2000US-0225266.
XX      14-AUG-2000; 2000US-0225267.
XX      14-AUG-2000; 2000US-0225268.
XX      14-AUG-2000; 2000US-0225270.
XX      14-AUG-2000; 2000US-022547.
XX      14-AUG-2000; 2000US-0225757.
XX      14-AUG-2000; 2000US-0225758.
XX      14-AUG-2000; 2000US-0225759.
XX      18-AUG-2000; 2000US-0226279.
XX      22-AUG-2000; 2000US-0226681.
XX      22-AUG-2000; 2000US-0226868.
XX      22-AUG-2000; 2000US-0227182.
XX      23-AUG-2000; 2000US-0227009.
XX      30-AUG-2000; 2000US-0228924.
XX      01-SEP-2000; 2000US-0229287.
XX      01-SEP-2000; 2000US-0229343.
XX      01-SEP-2000; 2000US-0229344.
XX      01-SEP-2000; 2000US-0229345.
XX      05-SEP-2000; 2000US-0229509.
XX      05-SEP-2000; 2000US-0229513.
XX      06-SEP-2000; 2000US-0230437.
XX      06-SEP-2000; 2000US-0230438.
XX      08-SEP-2000; 2000US-0231242.
XX      08-SEP-2000; 2000US-0231243.
XX      08-SEP-2000; 2000US-0231244.
XX      08-SEP-2000; 2000US-0231413.
XX      08-SEP-2000; 2000US-0231414.
XX      08-SEP-2000; 2000US-0232080.
XX      08-SEP-2000; 2000US-0232081.
XX      12-SEP-2000; 2000US-0231968.
XX      14-SEP-2000; 2000US-0232397.
XX      14-SEP-2000; 2000US-0232398.
XX      14-SEP-2000; 2000US-0232399.
XX      14-SEP-2000; 2000US-0232400.
XX      14-SEP-2000; 2000US-0233063.
XX      14-SEP-2000; 2000US-0233064.
XX      14-SEP-2000; 2000US-0233065.
XX      21-SEP-2000; 2000US-0234223.
XX      21-SEP-2000; 2000US-0234274.
XX
```

PR 25-SEP-2000; 2000US-0234997.  
PR 26-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.

PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465570/50.  
XX  
PT Isolated nucleic acid molecule encoding a reproductive system antigen -  
PT is used in preventing, treating or ameliorating a medical condition -  
XX  
XX Disclosure; SEQ ID NO 7825; 1297pp + Sequence Listing; English.  
XX  
XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention.  
XX  
SQ Sequence 2295 BP; 601 A; 537 C; 575 G; 582 T; 0 other;  
  
Query Match 75.6%; Score 31; DB 22; Length 2295;  
Best Local Similarity 82.9%; Pred. No. 0.019;  
Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
  
QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCCA 41  
||| ||||| ||||| ||||| ||| |||||  
Db 684 CCGTCTCTACTAAATAACAAAAATTAGCCTGGTGCCA 644  
  
RESULT 8  
AAL37445  
ID AAL37445 standard; DNA; 31730 BP.  
XX  
AC AAL37445;  
XX  
DT 08-JAN-2002 (first entry)  
XX  
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3810.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155367-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01338.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180528.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.

PR	14-AUG-2000;	2000US-0224518.
PR	14-AUG-2000;	2000US-0224519.
PR	14-AUG-2000;	2000US-0225213.
PR	14-AUG-2000;	2000US-0225214.
PR	14-AUG-2000;	2000US-0225266.
PR	14-AUG-2000;	2000US-0225267.
PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226868.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229509.
PR	05-SEP-2000;	2000US-0229513.
PR	06-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0233297.
PR	14-SEP-2000;	2000US-0233298.
PR	14-SEP-2000;	2000US-0233399.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0232401.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233064.
PR	14-SEP-2000;	2000US-0233065.
PR	21-SEP-2000;	2000US-0234223.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0241826.
PR	01-NOV-2000;	2000US-0244617.
PR	08-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246609.
PR	08-NOV-2000;	2000US-0246610.
PR	08-NOV-2000;	2000US-0246611.
PR	08-NOV-2000;	2000US-0246613.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
XX	05-JAN-2001;	2001US-0259678.
XX	(HUMA-)	HUMAN GENOME SCI INC.
PI	Rosen CA,	Barash SC, Ruben SM;
XX	WPI;	2001-451937/48.
DR	Isolated polypeptide for treating, preventing and/ or prognosing disorders related to the musculoskeletal system including musculoskeletal cancers and also for testing and detection e.g. diagnosis -	
XX	Example 2;	SEQ ID NO 3810; 781pp + Sequence Listing; English.
CC	The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABR03087-ABR04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;	



CC and (f) infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 31730 BP; 10793 A; 6331 C; 5682 G; 8924 T; 0 other;  
Query Match 75.6%; Score 31; DB 22; Length 31730;  
Best Local Similarity 82.9%; Pred. NO. 0.031;  
Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 41  
II IIIIIIIII IIIIIIIIIII III III IIIIIII  
Db 8860 CCATCTCTACTAAATAACAAAAACACGCGCGTGCGCA 8900  
RESULT 9  
ABQ88164  
ID ABQ88164 standard; cDNA; 86080 BP.  
XX  
AC ABQ88164;  
XX  
AC  
DT 18-SEP-2002 (first entry)  
XX  
DE Human osteoblast differentiation related cDNA SEQ ID NO 71.  
XX  
KW Human; osteoblast; stem cell differentiation; bone tissue deposition;  
KW osteoporosis; osteopathic; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200250301-A2.  
XX  
PD 27-JUN-2002.  
XX  
PF 18-DEC-2001; 2001WO-US48276.  
XX  
PR 18-DEC-2000; 2000US-255882P.  
PR 24-APR-2001; 2001US-285691P.  
XX  
XX  
PA (GENE-) GENE LOGIC INC.  
PA (PROC ) PROCTER & GAMBLE CO.  
XX  
XX  
PI JI D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;  
PI Mertz L;  
XX  
XX  
WPI; 2002-557663/59.  
XX  
XX  
PS Use of genes and their expression profiles associated with osteoblast  
PT differentiation for screening modulators bone formation, for diagnosing  
PT or treating e.g. osteoporosis, or as markers for the differentiation  
PT process -  
XX  
XX  
PS Claim 1; SEQ ID NO 71; 78pp + Sequence Listing; English.  
XX  
XX  
CC The invention relates to genes and their expression profiles are used  
CC for:  
CC (a) screening modulators of precursor stem cell differentiation into  
CC osteoblasts, or bone tissue deposition;  
CC (b) diagnosing abnormal deposition of bone tissue, abnormal rate of  
CC osteoblast formation or osteoporosis; or  
CC (c) treating or monitoring treatment of the conditions cited in (b), or  
CC monitoring the progression of bone tissue deposition.  
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid  
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy,  
CC drug-induced abnormalities in bone formation or bone loss, conditions  
CC that involve altered bone metabolism (e.g. idiopathic juvenile  
CC osteoporosis), skeletal disease linked to breast cancer, mastocytosis,  
CC Fanconi syndrome or fibrous dysplasia. The present sequence is that of an  
CC osteoblast differentiation associated cDNA marker of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 86080 BP; 21583 A; 21475 C; 21694 G; 21328 T; 0 other;  
Query Match 75.6%; Score 31; DB 24; Length 86080;  
Best Local Similarity 82.9%; Pred. NO. 0.037;  
Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 41  
II IIIIIIIII IIIIIIIIIII III III IIIIIII  
Db 38218 CCATCTCTACTAAATAACAATTAGCTGGCGGTGGCGCA 38258  
RESULT 10  
ABK83561  
ID ABK83561 standard; cDNA; 86080 BP.  
XX  
AC ABK83561;  
XX  
AC  
DT 14-AUG-2002 (first entry)  
XX  
DE Human cDNA differentially expressed in granulocytic cells #132.  
XX  
KW Human; ss; granulocytic cell; DNA chip; bacterial infection;  
KW viral infection; parasitic infection; protozoal infection;  
KW fungal infection; sterile inflammatory disease; psoriasis;  
KW rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;  
KW cardiac reperfusion injury; renal reperfusion injury; ARDS;  
KW adult respiratory distress syndrome; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; periodontal disease;  
KW granulocyte activation; chronic inflammation; allergy.  
XX  
OS Homo sapiens.  
XX  
PN WO200228999-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 03-OCT-2001; 2001WO-US30821.  
XX  
PR 03-OCT-2000; 2000US-237189P.  
XX  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
XX  
PI Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;  
PI WPI; 2002-435328/46.  
XX  
XX  
PS Detecting granulocyte activation by detecting differential expression  
PT of genes associated with granulocyte activation, which serves as  
PT diagnostic markers that is useful for monitoring disease states and  
PT drug toxicity -  
XX  
XX  
PS Claim 1; SEQ ID NO 132; 114pp; English.  
XX  
XX  
CC The invention relates to detecting (M1) granulocyte (GC) activation  
CC (GCA), by detecting the level of expression of gene(s) (Gs) identified by  
CC DNA chip analysis as given in the specification, and comparing  
CC the expression level to an expression level in an unactivated  
CC GC, where differential expression of Gs is indicative of GCA.  
CC Also included are modulating (M2) GA by contacting GC with an agent  
CC that alters the expression of at least one gene in Gs; (2) screening (M3)  
CC for an agent capable of modulating GCA or an inflammation (especially  
CC chronic) in a tissue, an allergic response in a subject, exposure of a  
CC subject to a pathogen or sterile inflammatory disease using the  
CC gene expression profile; (3) detecting (M4) an inflammation (especially  
CC chronic) in a tissue, an allergic response in a subject, exposure of a  
CC subject to a pathogen or sterile inflammatory disease, by detecting the  
CC level of expression in a sample of the tissue of gene(s) from Gs, where  
CC the level of expression of the gene is indicative of inflammation;  
CC (4) treating (M5) an inflammation (especially chronic) or in a tissue,  
CC an allergic response in a subject, exposure of a subject to a pathogen  
CC or sterile inflammatory disease, by contacting a tissue having

4  
0  
4  
0  
0  
0  
**1**  
0  
0  
0  
0  
0  
0  
4  
[  
]  
0  
~  
1  
4  
4  
0

AA	
AC	ABL67239;
XX	
DT	15-MAY-2002 (first entry)
XX	



PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 05-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
(HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Barash SC, Ruben SM;

PI WPI; 2001-465570/50.

DR Isolated nucleic acid molecule encoding a reproductive system antigen -  
PT is used in preventing, treating or ameliorating a medical condition -

PS Disclosure; SEQ ID NO 8680; 1297pp + Sequence Listing; English.

XX The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention.

SQ Sequence 264 BP; 61 A; 68 C; 92 G; 43 T; 0 other;

Query Match 74.6%; Score 30.6; DB 22; Length 264;

Best Local Similarity 84.6%; Pred. No. 0.018;

Matches 33; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGAGTGTGG 39

Db 105 CCCGTCTCTACTAAAAATACAAAAAGCTGGCGTGTGG 143

RESULT 14

ABL98556

ID ABL98556 standard; DNA; 264 BP.

XX ABL98556;

XX

DT	21-JUN-2002 (first entry)	
XX	Human testicular antigen encoding DNA fragment SEQ ID NO: 3208.	PR 14-SEP-2000; 2000US-0232401.
DE		PR 14-SEP-2000; 2000US-0233063.
XX		PR 14-SEP-2000; 2000US-0233064.
KW	Human; testicular antigen; testes; cancer; metastasis; immune disorder;	PR 14-SEP-2000; 2000US-0233065.
KW	reproductive system disorder; urinary system disorder; gene therapy;	PR 21-SEP-2000; 2000US-0234223.
KW	cardiovascular disorder; respiratory disorder; neurological disorder;	PR 25-SEP-2000; 2000US-0234274.
KW	gastrointestinal disease; infection; cytostatic; gene; ds.	PR 25-SEP-2000; 2000US-0234998.
XX		PR 26-SEP-2000; 2000US-0235484.
OS	Homo sapiens.	PR 27-SEP-2000; 2000US-0235834.
XX		PR 27-SEP-2000; 2000US-0235836.
PN	WO200155317-A2.	PR 29-SEP-2000; 2000US-0236327.
XX		PR 29-SEP-2000; 2000US-0236367.
PD	02-AUG-2001.	PR 29-SEP-2000; 2000US-0236368.
XX		PR 29-SEP-2000; 2000US-0236369.
PF	17-JAN-2001; 2001WO-US01329.	PR 29-SEP-2000; 2000US-0236370.
XX		PR 02-OCT-2000; 2000US-0236802.
XX		PR 02-OCT-2000; 2000US-0237037.
PR	31-JAN-2000; 2000US-0179065.	PR 02-OCT-2000; 2000US-0237038.
PR	04-FEB-2000; 2000US-0180628.	PR 02-OCT-2000; 2000US-0237039.
PR	24-FEB-2000; 2000US-0184664.	PR 02-OCT-2000; 2000US-0237039.
PR	02-MAR-2000; 2000US-0186350.	PR 02-OCT-2000; 2000US-0237040.
PR	16-MAR-2000; 2000US-0189874.	PR 13-OCT-2000; 2000US-0239935.
PR	17-MAR-2000; 2000US-0190076.	PR 13-OCT-2000; 2000US-0239937.
PR	18-APR-2000; 2000US-0198123.	PR 20-OCT-2000; 2000US-0240960.
PR	19-MAY-2000; 2000US-0205515.	PR 20-OCT-2000; 2000US-0241221.
PR	07-JUN-2000; 2000US-0209467.	PR 20-OCT-2000; 2000US-0241785.
PR	28-JUN-2000; 2000US-0214886.	PR 20-OCT-2000; 2000US-0241786.
PR	30-JUN-2000; 2000US-0215135.	PR 20-OCT-2000; 2000US-0241787.
PR	07-JUL-2000; 2000US-0216647.	PR 20-OCT-2000; 2000US-0241808.
PR	07-JUL-2000; 2000US-0216880.	PR 20-OCT-2000; 2000US-0241809.
PR	11-JUL-2000; 2000US-0217487.	PR 20-OCT-2000; 2000US-0241826.
PR	11-JUL-2000; 2000US-0217496.	PR 01-NOV-2000; 2000US-0244617.
PR	14-JUL-2000; 2000US-0218290.	PR 08-NOV-2000; 2000US-0246474.
PR	26-JUL-2000; 2000US-0220963.	PR 08-NOV-2000; 2000US-0246475.
PR	26-JUL-2000; 2000US-0220964.	PR 08-NOV-2000; 2000US-0246476.
PR	14-AUG-2000; 2000US-0224518.	PR 08-NOV-2000; 2000US-0246477.
PR	14-AUG-2000; 2000US-0224519.	PR 08-NOV-2000; 2000US-0246478.
PR	14-AUG-2000; 2000US-0225213.	PR 08-NOV-2000; 2000US-0246523.
PR	14-AUG-2000; 2000US-0225214.	PR 08-NOV-2000; 2000US-0246524.
PR	14-AUG-2000; 2000US-0225266.	PR 08-NOV-2000; 2000US-0246525.
PR	14-AUG-2000; 2000US-0225267.	PR 08-NOV-2000; 2000US-0246526.
PR	14-AUG-2000; 2000US-0225268.	PR 08-NOV-2000; 2000US-0246527.
PR	14-AUG-2000; 2000US-0225270.	PR 08-NOV-2000; 2000US-0246528.
PR	14-AUG-2000; 2000US-0225447.	PR 08-NOV-2000; 2000US-0246532.
PR	14-AUG-2000; 2000US-0225757.	PR 08-NOV-2000; 2000US-0246609.
PR	14-AUG-2000; 2000US-0225758.	PR 08-NOV-2000; 2000US-0246610.
PR	14-AUG-2000; 2000US-0225759.	PR 08-NOV-2000; 2000US-0246611.
PR	18-AUG-2000; 2000US-0226279.	PR 08-NOV-2000; 2000US-0246613.
PR	22-AUG-2000; 2000US-0226681.	PR 17-NOV-2000; 2000US-0249207.
PR	22-AUG-2000; 2000US-0226686.	PR 17-NOV-2000; 2000US-0249208.
PR	22-AUG-2000; 2000US-0227182.	PR 17-NOV-2000; 2000US-0249209.
PR	23-AUG-2000; 2000US-0227009.	PR 17-NOV-2000; 2000US-0249210.
PR	30-AUG-2000; 2000US-0228924.	PR 17-NOV-2000; 2000US-0249211.
PR	01-SEP-2000; 2000US-0229287.	PR 17-NOV-2000; 2000US-0249212.
PR	01-SEP-2000; 2000US-0229343.	PR 17-NOV-2000; 2000US-0249213.
PR	01-SEP-2000; 2000US-0229344.	PR 17-NOV-2000; 2000US-0249214.
PR	01-SEP-2000; 2000US-0229345.	PR 17-NOV-2000; 2000US-0249215.
PR	05-SEP-2000; 2000US-0229509.	PR 17-NOV-2000; 2000US-0249216.
PR	05-SEP-2000; 2000US-0229513.	PR 17-NOV-2000; 2000US-0249217.
PR	06-SEP-2000; 2000US-0230437.	PR 17-NOV-2000; 2000US-0249218.
PR	06-SEP-2000; 2000US-0230438.	PR 17-NOV-2000; 2000US-0249244.
PR	08-SEP-2000; 2000US-0231242.	PR 17-NOV-2000; 2000US-0249245.
PR	08-SEP-2000; 2000US-0231243.	PR 17-NOV-2000; 2000US-0249245.
PR	08-SEP-2000; 2000US-0231244.	PR 17-NOV-2000; 2000US-0249265.
PR	08-SEP-2000; 2000US-0231413.	PR 17-NOV-2000; 2000US-0249297.
PR	08-SEP-2000; 2000US-0231414.	PR 17-NOV-2000; 2000US-0249299.
PR	08-SEP-2000; 2000US-0232080.	PR 17-NOV-2000; 2000US-0249300.
PR	08-SEP-2000; 2000US-0232081.	PR 01-DEC-2000; 2000US-0250160.
PR	12-SEP-2000; 2000US-0231968.	PR 01-DEC-2000; 2000US-0250391.
PR	14-SEP-2000; 2000US-0232397.	PR 05-DEC-2000; 2000US-0251030.
PR	14-SEP-2000; 2000US-0232398.	PR 05-DEC-2000; 2000US-0251988.
PR	14-SEP-2000; 2000US-0232399.	PR 05-DEC-2000; 2000US-0256719.
PR	14-SEP-2000; 2000US-0232400.	PR 06-DEC-2000; 2000US-0251479.









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:51:09 ; Search time 118.293 Seconds  
(without alignments)  
5613.764 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_175\_215

Perfect score: 41

Sequence: 1 cctactctactgaaataatay.....aaaagctagactggtggca 41

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estnu:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estov:\*

7: em\_estro:\*

8: em\_estro:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estfun:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pin:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	40.6	99.0	382	14 T06700	T06700 EST04589 Fe
2	40.6	99.0	449	17 A0424894	A0424894 CITBT-EI-
3	39	95.1	364	17 A0428360	A0428360 CITBT-EI-
c 4	32.6	79.5	484	9 AA772119	AA772119 a140c06.s
5	31.6	77.1	679	17 AG175523	AG175523 Pan trogl
6	31.6	77.1	994	14 BQ653382	BQ653382 AGENCOURT

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	31.6	77.1	1103	14 BQ892326	BQ892326 AGENCOURT
8	31	75.6	319	17 A068839	A068839 HS_2255_A
c 9	31	75.6	420	14 T84104	T84104 Yd46f07.g1
10	31	75.6	445	17 A0154068	A0154068 HS_2236_B
c 11	31	75.6	468	9 A1638655	A1638655 tt25c03.x
c 12	31	75.6	487	14 R48796	R48796 YJ69f04.g1
c 13	31	75.6	588	12 BF213410	BF213410 601845167
c 14	31	75.6	578	17 A013280	A013280 RPT11-23
c 15	31	75.6	659	17 A048643	A048643 Pan trogl
c 16	31	75.6	909	14 BQ882472	BQ882472 AGENCOURT
c 17	31	75.6	1070	12 BG468620	BG468620 602510219
c 18	30.6	74.6	293	9 AA830965	AA830965 oc62404.s
c 19	30.6	74.6	339	9 AA483576	AA483576 nc75a06.s
c 20	30.6	74.6	346	12 BG432647	BG432647 602496318
c 21	30	73.2	266	9 AA533977	AA533977 nj95a07.s
c 22	30	73.2	369	10 AW962268	AW962268 EST374341
c 23	30	73.2	391	10 AV764272	AV764272 AV764272
c 24	30	73.2	396	17 A0261458	A0261458 CITBT-EI-
c 25	30	73.2	397	9 AA568331	AA568331 nf15c01.s
c 26	30	73.2	451	9 AA505035	AA505035 ab04906.r
c 27	30	73.2	519	17 A0357077	A0357077 CITBT-EI-
c 28	30	73.2	629	14 BQ417452	BQ417452 lk38b11.y
c 29	29.8	72.7	255	10 AW868364	AW868364 MRL-SN006
c 30	29.8	72.7	509	9 AL119907	AL119907 DRFZp7610
c 31	29.8	72.7	601	9 AL706735	AL706735 DRFZp686A
c 32	29.8	72.7	2943	11 AF318375	AF318375 Homo sap1
c 33	29.6	72.2	672	17 AG048010	AG048010 Pan trogl
c 34	29.6	72.2	679	17 AG144787	AG144787 Pan trogl
c 35	29.6	72.2	4553	17 AF152105	AF152105 AF152105
c 36	29.4	71.7	228	13 B1031811	B1031811 IL5-MT026
c 37	29.4	71.7	355	14 F08174	F08174 HSC2SC041.n
c 38	29.4	71.7	375	10 AV710483	AV710483 AV710483
c 39	29.4	71.7	380	14 T74665	T74665 YC57q03.r1
c 40	29.4	71.7	384	9 AA019829	AA019829 zg60a02.s
c 41	29.4	71.7	391	10 AV761983	AV761983 AV761983
c 42	29.4	71.7	414	14 W96435	W96435 ze43c03.r1
c 43	29.4	71.7	432	9 AA501407	AA501407 ne67a05.s
c 44	29.4	71.7	432	17 A0146914	A0146914 HS_2246_A
c 45	29.4	71.7	434	10 BE676916	BE676916 7d48a01.x

ALIGNMENTS

RESULT 1	T06700/c	T06700	382 bp	mRNA	linear	EST 30-JUN-1993
LOCUS	EST04589	Fetal brain, Stratagene (cat#936206)	Homo sapiens	CDNA		
DEFINITION	clone HFBDX16 similar to EST containing Alu repeat, mRNA sequence.					
ACCESSION	T06700					
VERSION	T06700.1	GI:317849				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 382)					
AUTHORS	Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.					
TITLE	3,400 expressed sequence tags identify diversity of transcripts from human brain					
JOURNAL	Nat. Genet. 4, 256-267 (1993)					
MEDLINE	93364420					
COMMENT	Contact: Adams, MD The Institute for Genomic Research 932 Clopper Road, Gaithersburg, MD 20878 Tel: 3018699056 Fax: 3018699423 Email: mdadams@tigr.org Seq primer: M13-21. Location/Qualifiers 1. .382 /organism="Homo sapiens" /db_xref="ATCC (inhost):83354"					
FEATURES	source					

```

/db_xref="taxon:9606"
/clone_lib="HFBDX16"
/clone_lib="Fetal brain, Stratagene (cat#936206)"
/notes="Vector: LambdaZAP-II; 17-18 wk gestation, female;
oligo-dT + random primed cDNA synthesis; lambdaZAP-II
vector, 1.0kb average inser size."
BASE COUNT      76 a   101 c   90 g   111 t   4 others
ORIGIN

Query Match      99.0%; Score 40.6; DB 14; Length 382;
Best Local Similarity 97.6%; Pred. No. 0.0015;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGGCA 41
|||||
Db 186 CCTATCTCTACTGAAATACAAAAAGCTAGACGTGGTGGCA 146
|||||

RESULT 2
AQ424894      449 bp   DNA   linear   GSS 24-MAR-1999
LOCUS
DEFINITION
CITBI-El-2576P11.TF CITBI-El Homo sapiens genomic clone 2576P11,
DNA sequence.
ACCESSION
AQ424894
VERSION
AQ424894.1 GI:4498160
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 449)
AUTHORS
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready
Map Building
JOURNAL
Unpublished (1997)
COMMENT
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
location/Qualifiers
1..364
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2576K5"
/clone_lib="CITBI-El"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelobAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
BASE COUNT      106 a   82 c   98 g   78 t
ORIGIN

Query Match      95.1%; Score 39; DB 17; Length 364;
Best Local Similarity 95.1%; Pred. No. 0.0052;
Matches 39; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTCAAAATAYAAAAAGCTAGACGTGGTGGCA 41
|||||
Db 172 CCTATCTCTACTCAAAATACATAAAGCTAGACGTGGTGGCA 212
|||||

RESULT 4
AA772119/c
LOCUS
DEFINITION
AA772119      484 bp   mRNA   linear   EST 31-DEC-1998
ai40c06.s1 Soares_parathyroid_tumor_NBHPA Homo sapiens cDNA clone
1359466 3', similar to contains Alu repetitive element;contains
element PTR7 repetitive element ;, mRNA sequence.
ACCESSION
AA772119
VERSION
AA772119.1 GI:2823902
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 484)
AUTHORS
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgaps-r@mail.nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.

```







```

RESULT 12
R48796/c
LOCUS
DEFINITION
    R48796          487 bp      mRNA      linear      EST 18-MAY-1995
    YJ69F04.s1 Soares breast 2NBHbst Homo sapiens cDNA clone
    IMAGE:154015 3' similar to contains Alu repetitive element; , mRNA
    sequence.
ACCESSION
R48796
VERSION
R48796.1      GI:810822
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 487)
AUTHORS
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,I., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
,R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
CONTACT: Wilton RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1188
High quality sequence stops: 315 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1188 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 315.
FEATURES
Location/Qualifiers
1. .487
/organism="Homo sapiens"
/db_xref="GDB:566398"
/db_xref="taxon:9606"
/clone_lib="IMAGE:154015"
/clone_lib="Soares breast 2NBHbst"
/sex="Female"
/dev_stage="adult"
/lab_host="PH10B (ampicillin resistant)"
/note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(GT) primer [5'
TGTTACCATCTCAAGTGGAGCGCGCCCTTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of a modified pT7T3 vector (Pharmacia).
Library went through one round of normalization to a Cot =
230. Library constructed by Bento Soares and M.Fatima
Bonaldo."
BASE COUNT      92 a  119 c  139 g  126 t  11 others
ORIGIN
Query Match      75.6%; Score 31; DB 14; Length 487;
Best Local Similarity 82.9%; Pred. No. 2.2;
Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 41
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 410 CCCATCTCTACTAAAATACAAAATCCAGCGGTGGTGCA 370

RESULT 13
BF213410
LOCUS
DEFINITION
    BF213410        568 bp      mRNA      linear      EST 06-NOV-2000
    601845167F1.NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4070393 5',
    mRNA sequence.
ACCESSION
BF213410
VERSION
BF213410.1    GI:11106996
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 578)
AUTHORS
Adams,M.D., Boursley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building (1998)
Unpublished (1998)
Other_GSSs: RPCI11-23B18.TVB RPCI11-23B18.TPB
CONTACT: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research

KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 568)
AUTHORS
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ARCC
CDNA library preparation: CLONETECH Laboratories, Inc.
CDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM914 row: e column: 18
High quality sequence stop: 554.
FEATURES
Location/Qualifiers
1. .568
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4070393"
/clone_lib="NIH_MGC_55"
/tissue_type="from acute myelogenous leukemia"
/lab_host="PH10B (T1 phage-resistant)"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggcgccctggcc); Site_2: SfiI (ggccattatggcc
); Double-stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCGCATTTATGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGCGCGGCACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT      154 a  120 c  147 g  147 t
ORIGIN
Query Match      75.6%; Score 31; DB 12; Length 568;
Best Local Similarity 82.9%; Pred. No. 2.2;
Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGCA 41
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 171 CCTGTCTCTACTAAAAATACAGAAATTAGCGGTGGTGCA 211

RESULT 14
AQ013280/c
LOCUS
DEFINITION
    AQ013280        578 bp      DNA      linear      GSS 14-APR-1999
    RPCI11-23B18.TKBR RPCI-11 Homo sapiens genomic clone RPCI-11-23B18,
    DNA sequence.
ACCESSION
AQ013280
VERSION
AQ013280.1    GI:3185845
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 578)
AUTHORS
Adams,M.D., Boursley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building (1998)
Unpublished (1998)
Other_GSSs: RPCI11-23B18.TVB RPCI11-23B18.TPB
CONTACT: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research

```

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html

Class: BAC ends.

# FEATURES

source

Location/Qualifiers

1. .578

/organism="Homo sapiens"

/db\_xref="GDB:7508489"

/db\_xref="taxon:9606"

/clone="RPCI-11-23B18"

/clone\_lib="RPCI-11"

/sex="Male"

/cell\_type="Lymphocytes"

/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;

RPCI11 Human Male BAC Library"

171 a 122 c 138 g 147 t

BASE COUNT

ORIGIN

Query Match 75.6%; Score 31; DB 17; Length 578;

Best Local Similarity 82.9%; Pred. No. 2.2;

Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGGCA 41

Db 340 CCTGCTCTCTACTAAAAATACAAAAAATAGGCGTAGTGGCA 300

## RESULT 15

AG048643/C

LOCUS

AG048643 Pan troglodytes DNA, clone: PTB-028I12.F, genomic survey sequence.

DEFINITION

AG048643

VERSION

AG048643.1 GI:16585535

KEYWORDS

SOURCE

Pan troglodytes male lymphoblast DNA, clone\_lib:PTB Chimpanzee Male

BAC Library clone:PTB-028I12.F.

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

1

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

BAC end sequences of Library PTB

Unpublished

2 (bases 1 to 659)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail: chimps@gscl.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,

Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end

was generated during the R&D process and may have higher chance of

clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI

Location/Qualifiers

1. .659

/organism="Pan troglodytes"

/db\_xref="taxon:9598"

## FEATURES

source

/clone="PTB-028I12.F"

/sex="male"

/cell\_type="lymphoblast"

/clone\_lib="PTB Chimpanzee Male BAC Library"

BASE COUNT 107 a 228 c 155 g 169 t

ORIGIN

Query Match 75.6%; Score 31; DB 17; Length 659;

Best Local Similarity 82.9%; Pred. No. 2.2;

Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGGCA 41

Db 384 CCTGCTCTCTACTAAAAATACAAAAAATCAGGCGTGGTGGCA 344

Search completed: February 11, 2003, 03:31:10

Job time : 122.283 secs









Matches	29;	Conservative	1;	Mismatches	5;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

QY 1 CCTATCTCTACTGAAATAAYAAAAAGCTAGACGTG 35  
 || ||||| |||||:|||| ||| ||||  
 Db 51462 CCCATCTCTACTAAAAATACAAAAAGTAGCCGTG 51496

## RESULT 7

```

US-09-218-363-3/c
; Sequence 3, Application US/09218363
; Patent No. 6387616
; GENERAL INFORMATION:
; APPLICANT: Ozelius, Laurie J.
; APPLICANT: Breakfield, Xandra O.
; TITLE OF INVENTION: TORSIN, TORSIN GENES, AND METHODS OF USE
; FILE REFERENCE: MGH-1184PA2
; CURRENT APPLICATION NUMBER: US/09/218,363
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 09/099,454
; EARLIER FILING DATE: 1998-06-18
; EARLIER APPLICATION NUMBER: 60/050,244
; EARLIER FILING DATE: 1997-06-19
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 3568

```

LENGTH: 3568

```

1  NAME: 9500
2  TYPE: DNA
3  ORGANISM: Homo sapien
4  FEATURE:
5  NAME/KEY: CDS
6  LOCATION: (94)...(1863)
7  NAME/KEY: misc_feature
8  LOCATION: (1)...(3568)
9  OTHER INFORMATION: n A,T,C or G

```

US-09-218-363-3

Query Match 62.9%; Score 25.8; DB 4; Length 3568;  
Best Local Similarity 76.9%; Pred. No. 0.22;  
Matches 30; Conservative 1; Mismatches 8; Indels 0

Qy 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTGGTGG 39  
 || ||||| |||||:||||| ||||| |||||  
 Db 396 CCGTGCTCTACTAAAAATACAAAAAATTAGATGGTGTGG 358

## RESULT 8

US-09-608-285A-42/c  
; Sequence 42, Application US/09608285A  
; Patent No. 6335013  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD33 LINE  
; TITLE OF INVENTION: POLYPEPTIDES

```

: FILE REFERENCE: 28110736570
:
: CURRENT APPLICATION NUMBER: US/09/608,285A
:
: CURRENT FILING DATE: 2000-06-30
:
: PRIOR APPLICATION NUMBER: 09/583,231
:
: PRIOR FILING DATE: 2000-05-26
:
: PRIOR APPLICATION NUMBER: 09/557,800
:
: PRIOR FILING DATE: 2000-04-25
:
: PRIOR APPLICATION NUMBER: 09/481,238
:
: PRIOR FILING DATE: 2000-01-11
:
: PRIOR APPLICATION NUMBER: 09/370,265
:
: PRIOR FILING DATE: 1999-08-09
:
: PRIOR APPLICATION NUMBER: PCT/US99/16180
:
: PRIOR FILING DATE: 1999-07-16
:
: PRIOR APPLICATION NUMBER: 09/350,836
:
: PRIOR FILING DATE: 1999-07-09
:
: PRIOR APPLICATION NUMBER: 09/273,447
:
: PRIOR FILING DATE: 1999-03-19
:
: PRIOR APPLICATION NUMBER: 09/244,444

```

```

; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 14747
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (13641)
; OTHER INFORMATION: n adenosine or guanine or cytosine or thymidine
US-09-608-285A-42

```

Query Match	62.9%	Score 25.8	DB 4	Length 14747
-------------	-------	------------	------	--------------

Query Match	62.9%	Score 25.8;	DB 4;	Length 14747;
Best Local Similarity	76.9%	Pred. No. 0.28;		
Matches 30; Conservative	1;	Mismatches 8;	Indels 0;	Cups 0;

**Qy** 1 CCTATCTCTACTGAANTAYAAAAAGCTAGACGTGGTGG 39  
|| ||||||| ||||||| :||| | | |||||  
**Dd** 3310 CCATCTCTACTAAATAATACAAAAAATTAGCATGTTGG 3272

## RESULT 9

```

RESOL. 3
US-09-608-285A-59/c
; Sequence 59, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES

```

FILE OF INVENTION: POLYPEP  
FILE REFERENCE: 28110/36570

```

> FILE REFERENCE: 28110/06370
> CURRENT APPLICATION NUMBER: US/09/608, 2854
> CURRENT FILING DATE: 2000-06-30
> PRIOR APPLICATION NUMBER: 09/583,231
> PRIOR FILING DATE: 2000-05-26
> PRIOR APPLICATION NUMBER: 09/557,800
> PRIOR FILING DATE: 2000-04-25
> PRIOR APPLICATION NUMBER: 09/481,238
> PRIOR FILING DATE: 2000-01-11
> PRIOR APPLICATION NUMBER: 09/370,265
> PRIOR FILING DATE: 1999-08-09
> PRIOR APPLICATION NUMBER: PCT/US99/16180
> PRIOR FILING DATE: 1999-07-16
> PRIOR APPLICATION NUMBER: 09/350,836
> PRIOR FILING DATE: 1999-07-09
> PRIOR APPLICATION NUMBER: 09/273,447
> PRIOR FILING DATE: 1999-03-19
> PRIOR APPLICATION NUMBER: 09/244,444
> PRIOR FILING DATE: 1999-02-04
> PRIOR APPLICATION NUMBER: 09/122,449
> PRIOR FILING DATE: 1998-07-24
> PRIOR APPLICATION NUMBER: 09/118,205
> PRIOR FILING DATE: 1998-07-16
> NUMBER OF SEQ ID NOS: 60
> SOFTWARE: Patent Ver. 2.0
> SEQ ID NO 59
> LENGTH: 15977
> TYPE: DNA
> ORGANISM: Homo sapiens
> FEATURE:
> OTHER INFORMATION: CD39-L4/L66 Gene Sequence
> NAME/KEY: CDS
> LOCATION: (245)..(461)
> NAME/KEY: CDS
> LOCATION: (1454)..(1533)
> NAME/KEY: CDS
> LOCATION: (2734)..(2877)

```

```
; NAME/KEY: CDS
; LOCATION: (4364)..(4439)
; NAME/KEY: CDS
; LOCATION: (4679)..(4714)
; NAME/KEY: CDS
; LOCATION: (5326)..(5414)
; NAME/KEY: CDS
; LOCATION: (5723)..(5802)
; NAME/KEY: CDS
; LOCATION: (6751)..(6812)
; NAME/KEY: CDS
; LOCATION: (7758)..(7859)
; NAME/KEY: CDS
; LOCATION: (8712)..(8852)
; NAME/KEY: CDS
; LOCATION: (9831)..(9887)
; NAME/KEY: CDS
; LOCATION: (11613)..(11728)
; NAME/KEY: CDS
; LOCATION: (13146)..(13691)
; NAME/KEY: CDS
; LOCATION: (15702)..(15839)
; NAME/KEY: misc_feature
; LOCATION: (14871)
; OTHER INFORMATION: n = a or c or g or t
US-09-608-285A-59

Query Match          62.9%; Score 25.8; DB 4; Length 15977;
Best Local Similarity 76.9%; Pred. No. 0.29;
Matches 30; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGCGTGGTGG 39
||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3310 CCATCTCTACTAAATAACAAAAAATTAGCATGGTGG 3272

RESULT 10
US-09-128-155-16/c
; Sequence 16, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 152331
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(152331)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

Query Match          62.9%; Score 25.8; DB 3; Length 152331;
Best Local Similarity 76.9%; Pred. No. 0.44;
Matches 30; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGCGTGGTGG 39
||| ||||| ||||| ||||| ||||| ||||| |||||
Db 107781 CCATCTCTACTAAATAACAAAAAATTAGCCAGTGG 107743

RESULT 11
```

```
US-09-851-896-3
; Sequence 3, Application US/09851896
; Patent No. 6410325
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEP
; FILE REFERENCE: RFS-0220
; CURRENT APPLICATION NUMBER: US/09/851,896
; CURRENT FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 70000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-851-896-3

Query Match          61.5%; Score 25.2; DB 4; Length 70000;
Best Local Similarity 84.4%; Pred. No. 0.66;
Matches 27; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGAC 32
||| ||||| ||||| ||||| ||||| ||||| |||||
Db 29288 CCTGTCTCTACTGAAATACAAAAAATTAGCC 29319

RESULT 12
US-09-385-982-12/c
; Sequence 12, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 581
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(581)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-12

Query Match          61.0%; Score 25; DB 4; Length 581;
Best Local Similarity 80.0%; Pred. No. 0.32;
Matches 28; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGTG 35
||| ||||| ||||| ||||| ||||| ||||| |||||
Db 188 CCTGTCTCTACTGAAATACAAAAAATTAGCCGGG 154

RESULT 13
US-09-641-638-376/c
; Sequence 376, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
```

```
; APPLICANT: Bouqueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 376
; LENGTH: 1001
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-878-153 : polymorphic base C or T
; NAME/KEY: misc.binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-878-153.mis1, potential
; NAME/KEY: misc.binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-878-153.mis2, potential complement
; NAME/KEY: primer.bind
; LOCATION: 349..369
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer.bind
; LOCATION: 839..859
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc.binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-878-153 potential probe
; OTHER INFORMATION: 12-878-153 potential probe
US-09-641-638-376

Query Match 61.0%; Score 25; DB 4; Length 1001;
Best Local Similarity 80.0%; Pred. No. 0.35;
Matches 28; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGCTG 35
Db 85 CCCATCTCTACTAAATAACAAAAAATTAGCGGAG 51

RESULT 14
US-09-701-685-1
; Sequence 1, Application US/09701685
; Patent No. 6387629
; GENERAL INFORMATION:
; APPLICANT: Schneider, Patrick
; APPLICANT: Yamamoto, Karen K.
; APPLICANT: French, Cynthia K.
; APPLICANT: Reprogen, Inc.
; TITLE OF INVENTION: Use of Cathepsin S in the Diagnosis and Treatment of
; TITLE OF INVENTION: Endometriosis
; FILE REFERENCE: 018002-001310US
; CURRENT APPLICATION NUMBER: US/09/701,685
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: WO PCT/US99/12335
; PRIOR FILING DATE: 1999-06-03
; PRIOR APPLICATION NUMBER: US 60/088,017
; PRIOR FILING DATE: 1998-06-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
```

```
; LENGTH: 1643
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(1065)
; OTHER INFORMATION: human cathepsin S
US-09-701-685-1

Query Match 61.0%; Score 25; DB 4; Length 1643;
Best Local Similarity 80.0%; Pred. No. 0.39;
Matches 28; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGCTG 35
Db 1311 CCCATCTCTACTAAATAACAAAAAATTAGCGGAG 1345

RESULT 15
US-09-014-969-1/c
; Sequence 1, Application US/09014969
; Patent No. 5965397
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/014,969
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2509 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-014-969-1

Query Match 61.0%; Score 25; DB 2; Length 2509;
Best Local Similarity 80.0%; Pred. No. 0.42;
Matches 28; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCTATCTCTACTGAAATAYAAAAAGCTAGACGCTG 35
Db 1730 CCTGCTCTACTAAATAACAAAAAATTAGCGGG 1696
```

Search completed: February 11, 2003, 05:42:25  
Job time : 55.1156 secs

---

Result No.	Score	Query			DB	ID	Description
		Match	Length				
C 1	31.6	77.1	9704	12	US-10-109-860-3	Sequence 3, Appli	
C 2	31	75.6	13730	10	US-09-764-877-3810	Sequence 3810, Ap	
C 3	31	75.6	167343	10	US-09-962-436-281	Sequence 281, App	
C 4	31	75.6	167343	10	US-09-964-824A-273	Sequence 273, App	
C 5	29.4	71.7	7032	9	US-09-974-238-124	Sequence 124, App	
C 6	29.4	71.7	10514	10	US-09-764-877-3470	Sequence 3470, Ap	
C 7	29.4	71.7	88191	10	US-09-799-799-3	Sequence 3, Appli	
C 8	27.8	67.8	3941	10	US-09-764-869-2372	Sequence 2372, Ap	
C 9	27.8	67.8	12822	10	US-09-764-847-1579	Sequence 1579, Ap	
C 10	27.8	67.8	51719	10	US-09-918-686-2	Sequence 2, Appli	
C 11	27.8	67.8	67998	10	US-09-880-107-3949	Sequence 3949, Ap	
C 12	27.8	67.8	92139	10	US-09-918-686-1	Sequence 1, Appli	
C 13	27.4	66.8	18860	10	US-09-764-877-2317	Sequence 2317, Ap	
C 14	27.4	66.8	45839	12	US-10-025-187-3	Sequence 3, Appli	
C 15	27.4	66.8	84539	10	US-09-962-436-36	Sequence 36, Appli	
C 16	26.8	65.4	4359	10	US-09-764-864-1717	Sequence 1717, Ap	
C 17	26.6	64.9	461	10	US-09-867-701-8502	Sequence 8502, Ap	
C 18	26.6	64.9	1066	10	US-09-764-877-2730	Sequence 2730, Ap	
C 19	26.6	64.9	2092	10	US-09-764-869-1930	Sequence 1930, Ap	

```

Query Match          75.6%; Score 31; DB 10; Length 167343;
Best Local Similarity 82.9%; Pred. No. 0.045; 6; Indels 0; Gaps 0;
Matches 34; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      1 CCTATCTCTACTGAAATAAYAAAAAGCTAGACGTGGTGCCA 41
       ||| ||||||| ||||||| ||||| ||| ||| ||||||
Db 134985 CCGTCTCTACTAAAATACAAAAAATTAGCGGTAGTGCCA 134945


RESULT 5
US-09-974-298-124
; Sequence 124, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 124
; LENGTH: 7032
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incyte ID No. US20020156263A1 474916.2
; NAME/KEY: unsure
; LOCATION: 3880-3917
; OTHER INFORMATION: a, t, c, g, or other
US-09-974-298-124


Query Match          71.7%; Score 29.4; DB 9; Length 7032;
Best Local Similarity 80.5%; Pred. No. 0.11; 7; Indels 0; Gaps 0;
Matches 33; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      1 CCTATCTCTACTGAAATAAYAAAAAGCTAGACGTGGTGCCA 41
       ||| ||||||| ||||||| ||||| ||| ||| ||||||
Db 2702 CCCATCTCTACTAAAATACAAAAAATTAGGTGTAGTGCCA 2742


RESULT 6
US-09-764-877-3470
; Sequence 3470, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3470
; LENGTH: 10514
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3470


Query Match          71.7%; Score 29.4; DB 10; Length 10514;
Best Local Similarity 80.5%; Pred. No. 0.12; 7; Indels 0; Gaps 0;
Matches 33; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      1 CCTATCTCTACTGAAATAAYAAAAAGCTAGACGTGGTGCCA 41
       ||| ||||||| ||||||| ||||| ||| ||| ||||||
Db 6181 CCCATCTCTACTAAAATACAAAAAATTAGGTGTAGTGCCA 6221

```



```

; SEQ ID NO 1579
; LENGTH: 12822
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1579

Query Match      67.8%; Score 27.8; DB 10; Length 12822;
Best Local Similarity 78.0%; Pred. No. 0.45;
Matches 32; Conservative 1; Mismatches 8; Indels 0; Gaps 0

QY      1 CCTATCTCTACTGAATATAYAAAAAGCTAGACGTGGTGCCA 41
Db      10062 CCTGCTCTGCTAAAAATACAAAAAACACGAGTGTGGTGA 10102
          ||| ||||| || ||||| ||||| || || ||||| ||
          ||| ||||| || ||||| ||||| || || ||||| ||

RESULT 10
US-09-918-686-2/c
; Sequence 2, Application US/09918686
; Patent No. US20020076720A1
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Prohl, Sean
; APPLICANT: Paepker, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; TITLE OF INVENTION: GENOMIC DELETIONS
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918.686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 51719
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1246, 2572, 2604
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-2

Query Match      67.8%; Score 27.8; DB 10; Length 51719;
Best Local Similarity 78.0%; Pred. No. 0.53;
Matches 32; Conservative 1; Mismatches 8; Indels 0; Gaps 0

QY      1 CCTATCTCTACTGAATATAYAAAAAGCTAGACGTGGTGCCA 41
Db      25322 CCCATCTCTACTAAAAATACAAATTAGCCGGGCATGGTGCCA 25482
          ||| ||||| ||||| ||||| || || ||||| ||
          ||| ||||| ||||| ||||| || || ||||| ||

RESULT 11
US-09-880-107-3949
; Sequence 3949, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880.107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3949
; LENGTH: 76798
; TYPE: DNA
; ORGANISM: Homo sapiens

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.  
  
OM nucleic - nucleic search, using sw model  
  
Run on: February 10, 2003, 20:43:59 ; Search time 16.6596 Seconds  
(without alignments)  
5542.256 Million cell updates/sec  
  
Title: US-09-942-310-2\_copy\_920\_960  
Perfect score: 41  
Sequence: 1 ctttctgtgggtgatttct.....crtgtgtaatcgtgcctctg 41  
  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
  
Searched: 2185239 seqs, 1125999159 residues  
  
Total number of hits satisfying chosen parameters: 4370478  
  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :				N_Geneseq_101002:*	
1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*				
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*				
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*				
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*				
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*				
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*				
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*				
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*				
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*				
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*				
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*				
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*				
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*				
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*				
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*				
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*				
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*				
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*				
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*				
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*				
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*				
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*				
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*				
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				Description	
Result No.	Score	Query Match	Length DB ID		
1	40.6	99.0	1680	AAD34214	Human CYP2D6 gene
2	40.6	99.0	6472	ABQ72215	Human CYP2D6 gene,
3	40.6	99.0	6472	ABQ72364	Human CYP2D6 gene,
4	40.6	99.0	9432	AAD34213	Human cytochrome p
5	40.2	98.0	1669	AAH26169	Human cytochrome p
6	40.2	98.0	1669	AAH26179	Human cytochrome p
7	31.6	77.1	5884	ABK3958	Human chemically p
8	31.6	77.1	5884	ABL32556	Human immune syste
9	23.6	57.6	2299	ABK36040	CDNA sequence #431

C 10	23.6	57.6	5884	24	ABK39959	Human chemically p
C 11	23.6	57.6	5884	24	ABL32557	Human immune syste
C 12	23.6	57.6	6281	22	ABA17450	Human nervous syst
C 13	23.6	57.6	6286	22	ABA17449	Human nervous syst
C 14	23.4	57.1	160755	23	AAH88704	Human DNA sequence
C 15	23	56.1	500	24	ABK62119	Rat sequence diffe
C 16	23	56.1	2405	23	ABL28472	Drosophila melanog
C 17	23	56.1	149671	24	ABK84797	Human cDNA differe
C 18	22.8	55.6	144460	21	AAZ93815	Olfactory receptor
C 19	22.6	55.1	285	19	AAV48130	z-chromosomal micr
C 20	22.6	55.1	800	24	ABQ89221	Human prostate exp
C 21	22.6	55.1	2494	23	ABL17592	Drosophila melanog
C 22	22.6	55.1	11273	23	ABL06992	Drosophila melanog
C 23	22.6	55.1	11530	23	ABL20948	Oligonucleotide fo
C 24	22.2	54.1	740	24	ABQ15472	Oligonucleotide fo
C 25	22.2	54.1	740	24	ABQ15473	Oligonucleotide fo
C 26	22.2	54.1	4629	23	ABL27094	Drosophila melanog
C 27	22	53.7	326	22	AAI88442	Human polynucleoti
C 28	22	53.7	584	24	ABQ27786	Oligonucleotide fo
C 29	22	53.7	584	24	ABQ27787	Oligonucleotide fo
C 30	22	53.7	1735	21	AAC60025	Human secreted pro
C 31	22	53.7	3318	20	AAZ28300	Rat neuronal immed
C 32	22	53.7	3521	21	AAZ57791	5' upstream and pr
C 33	22	53.7	5152	24	ABL32506	Human immune syste
C 34	22	53.7	5629	23	ABK42274	Genomic sequence #
C 35	22	53.7	6265	23	ABK42274	Genomic sequence #
C 36	22	53.7	6971	24	ABL33236	Human immune syste
C 37	22	53.7	8088	23	ABL29030	Drosophila melanog
C 38	22	53.7	8423	22	ABA15397	Human nervous syst
C 39	22	53.7	11089	22	AAK79653	Human immune/haema
C 40	22	53.7	11622	24	ABL32675	Human immune syste
C 41	22	53.7	13819	22	ABA15398	Human nervous syst
C 42	22	53.7	15944	22	ABA15399	Human nervous syst
C 43	22	53.7	16006	23	ABL14238	Drosophila melanog
C 44	22	53.7	16424	22	AAK68448	Human immune/haema
C 45	22	53.7	16424	22	AAK75677	Human immune/haema

ALIGNMENTS

RESULT 1  
AAD34214  
ID AAD34214 standard; DNA; 1680 BP.  
XX  
AC AAD34214;  
XX  
XX 16-JUL-2002 (first entry)  
DT  
XX Human CYP2D6 gene 5' flanking region containing polymorphic sites.  
DE  
DE Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; polymorphism; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT misc\_feature 36  
FT Location/Qualifiers  
FT /\*tag= a  
FT /note= "Polymorphic site"  
FT 194  
FT misc\_feature  
FT /\*tag= b  
FT /note= "Polymorphic site"  
FT 385  
FT misc\_feature  
FT /\*tag= c  
FT /note= "Polymorphic site"  
FT 620  
FT misc\_feature  
FT /\*tag= d  
FT /note= "Polymorphic site"  
FT 880  
FT misc\_feature  
FT /\*tag= e  
FT /note= "Polymorphic site"

```

FT misc_feature 942
FT /*tag= f
FT /*label= PS1
FT /*note= "Novel single nucleotide polymorphism (SNP);
FT given as R in the specification"
FT
FT misc_feature
FT 1255
FT /*tag= g
FT /*label= PS2
FT /*note= "Novel single nucleotide polymorphism (SNP);
FT given as Y in the specification"
FT
XX WO200218638-A2.
XX
XX PD 07-MAR-2002.
XX
XX XX 27-AUG-2001; 2001WO-IB01544.
XX PF
XX PR 30-AUG-2000; 2000GB-0021286.
XX
XX (GEMI-) GEMINI GENOMICS PLC.
XX
XX PI Risinger C, Andersson MK, Lewander T, Olliasson E;
XX
XX WPI; 2002-329785/36.
XX
XX New sequence determination oligonucleotides, useful for detecting
XX polymorphic sites in a 5' flanking region of a CYP2D6 gene, as
XX hybridization probes, as components of diagnostic assays, or in
XX ligase-based sequence determination -
XX
XX PS Claim 1; Fig 2; 63pp; English.
XX
XX The invention relates to sequence determination oligonucleotides for
XX detecting polymorphic sites in a 5' flanking region of cytochrome P450
XX 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many
XX different xenobiotics. Human CYP2D6 gene is located on chromosome 22.
XX The oligonucleotides may be used as in situ hybridisation probes, in
XX ligase-based sequence determination, as components of diagnostic assays,
XX as probes in sequence determination methods based on mismatches, as
XX hybridisation-based diagnostic assays, and as components of diagnostic
XX microarray. CYP2D6 is useful to predict variations in an individual's
XX ability to metabolise certain drugs. The present sequence is human
XX CYP2D6 gene 5' flanking region containing polymorphic sites.
XX
XX SQ Sequence 1680 BP; 413 A; 379 C; 539 G; 342 T; 7 other;
XX
XX Query Match 99.0%; Score 40.6; DB 24; Length 1680;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-06;
XX Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CTTGTGCGGTGATTTCTGCTGTGTAATCGTGCCCTG 41
XX |||||
XX Db 920 CTTGTGCGGTGATTTCTGCTGTGTAATCGTGCCCTG 960
XX
XX RESULT 2
XX ABQ72215
XX ID ABQ72215 standard; DNA; 6472 BP.
XX AC
XX ABQ72215;
XX
XX DT 02-SEP-2002 (first entry)
XX
XX DE Human CYP2D6 gene, SEQ ID NO:1 version #1.
XX
XX Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;
XX chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;
XX antiarrhythmic; arrhythmia; adrenoceptor antagonist; hypertension;
XX tricyclic antidepressant; procainamide; drug induced lupus syndrome;
XX environmentally linked disease; Parkinson's disease; haplotyping;
XX genotyping; haplotype; genetic variant; single nucleotide polymorphism;
XX SNP; drug screening; drug discovery; gene; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX variation replace (636, A)

```

```
FT      given as R in the specification; causes the
FT      amino acid substitution R88H"
FT      replace (1974, A)
FT      /*tag= q
FT      /label= PS13
FT      /note= "Known single nucleotide polymorphism (SNP);
FT      given as M in the specification; causes the
FT      amino acid substitution L91M"
FT      variation
FT      replace (1984, G)
FT      /*tag= r
FT      /label= PS14
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as R in the specification; causes the
FT      amino acid substitution H94R"
FT      variation
FT      replace (1997, G)
FT      /*tag= s
FT      /label= PS15
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as S in the specification"
FT      variation
FT      replace (2014, C)
FT      /*tag= t
FT      /label= PS16
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as Y in the specification; causes the
FT      amino acid substitution V104A"
FT      variation
FT      replace (2022, T)
FT      /*tag= u
FT      /label= PS17
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as W in the specification; together
FT      with PS18 causes the amino acid substitution
FT      T107F"
FT      variation
FT      replace (2023, T)
FT      /*tag= v
FT      /label= PS18
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as Y in the specification; together
FT      with PS17 causes the amino acid substitution
FT      T107F"
FT      variation
FT      replace (2028, G)
FT      /*tag= w
FT      /label= PS19
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as R in the specification; causes the
FT      amino acid substitution I109V"
FT      variation
FT      replace (2036, C)
FT      /*tag= x
FT      /label= PS20
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as Y in the specification"
FT      variation
FT      replace (2039, T)
FT      /*tag= y
FT      /label= PS21
FT      /note= "Known single nucleotide polymorphism (SNP);
FT      given as Y in the specification"
FT      intron
FT      2056..2605
FT      /*tag= z
FT      /number= 2
FT      /cons_splice= (5'site:NO, 3'site:YES)
FT      variation
FT      replace (2062, G)
FT      /*tag= aa
FT      /label= PS22
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as R in the specification"
FT      variation
FT      replace (2067, G)
FT      /*tag= ab
FT      /label= PS23
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as K in the specification"
FT      variation
FT      replace (2118, T)
FT      /*tag= ac
FT      /label= PS24
FT      /note= "Novel single nucleotide polymorphism (SNP);
```

```
FT      given as Y in the specification"
FT      replace (2170, A)
FT      /*tag= ad
FT      /label= PS25
FT      /note= "Known single nucleotide polymorphism (SNP);
FT      given as R in the specification"
FT      variation
FT      replace (2179, C)
FT      /*tag= ae
FT      /label= PS26
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as S in the specification"
FT      exon
FT      2606..2758
FT      /*tag= af
FT      /number= 3
FT      variation
FT      replace (2611, A)
FT      /*tag= ag
FT      /label= PS27
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as W in the specification; causes the
FT      amino acid substitution F120I"
FT      variation
FT      replace (2635, C)
FT      /*tag= ah
FT      /label= PS28
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as Y in the specification; causes the
FT      amino acid substitution W128R"
FT      variation
FT      replace (2659, A)
FT      /*tag= ai
FT      /label= PS29
FT      /note= "Novel single nucleotide polymorphism (SNP);
FT      given as R in the specification; together
FT      with PS30 causes the amino acid substitution
FT      V136I"
FT      variation
FT      replace (2661, C)

Query Match      99.0%; Score 40.6; DB 24; Length 6472;
Best Local Similarity 97.6%; Pred. No. 2.le-06;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 CTTTGTGGGTGATTTCCTGCTGTGAATCGTGCCCTG 41
        |||||
DB      301 CTTTGTGGGTGATTTCCTGCTGTGAATCGTGCCCTG 341

RESULT 3
ABQ72364
ID      ABQ72364 standard; DNA; 6472 BP.
XX      ABQ72364;
AC      ABQ72364;
XX      02-SEP-2002 (first entry)
DT      Human CYP2D6 gene, SEQ ID NO:1 version #2.
DE
XX      Human; cytochrome P450; subfamily IID polypeptide 6; CYP2D6; enzyme;
KW      chromosome 22q13.1; drug metabolism; detoxification; mono-oxygenase;
KW      antiarrhythmic; arrhythmia; adrenoreceptor antagonist; hypertension;
KW      tricyclic antidepressant; procainamide; drug induced lupus syndrome;
KW      environmentally linked disease; Parkinson's disease; haplotyping;
KW      genotyping; haplotype; genetic variant; single nucleotide polymorphism;
KW      SNP; drug screening; drug discovery; gene; ds.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      variation replace (636, A)
FT      /*tag= a
FT      /label= PS1
FT      /note= "Novel single nucleotide polymorphism (SNP)"
FT      variation replace (678, C)
FT      /*tag= b
FT      /label= PS2
FT      /note= "Novel single nucleotide polymorphism (SNP)"
```

```
FT variation replace (769, C)
FT /tag= c
FT /label= PS3
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (776, G)
FT /tag= d
FT /label= PS4
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (825, A)
FT /tag= e
FT /label= PS5
FT /note= "Known single nucleotide polymorphism (SNP)";
FT variation replace (915, C)
FT /tag= f
FT /label= PS6
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT CDS 1001..5217
FT /tag= g "cYP2D6"
FT /product=
FT 1001..1180
FT /tag= h
FT /number= 1
FT variation replace (1019, A)
FT /tag= i
FT /label= PS7
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V7M"
FT variation replace (1031, A)
FT /tag= j
FT /label= PS8
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution W11M"
FT variation replace (1100, T)
FT /tag= k
FT /label= PS9
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution P34S"
FT intron 1181..1883
FT /tag= l
FT /number= 1
FT variation replace (1827, C)
FT /tag= m
FT /label= PS10
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (1843, G)
FT /tag= n
FT /label= PS11
FT /note= "Known single nucleotide polymorphism (SNP)".
FT exon 1884..2055
FT /tag= o
FT /number= 2
FT variation replace (1966, A)
FT /tag= p
FT /label= PS12
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution R88H"
FT variation replace (1974, A)
FT /tag= q
FT /label= PS13
FT /note= "Known single nucleotide polymorphism (SNP);
FT causes the amino acid substitution L91M"
FT variation replace (1984, G)
FT /tag= r
FT /label= PS14
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution H94R"
FT variation replace (1997, G)
FT /tag= s
FT /label= PS15
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (2014, C)
FT /tag= t
FT /label= PS16
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution V104A"
FT variation replace (2022, T)
FT /tag= u
FT /label= PS17
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS18 causes the amino acid
FT substitution T107F"
FT variation replace (2023, T)
FT /tag= v
FT /label= PS18
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS17 causes the amino acid
FT substitution T107F"
FT variation replace (2028, G)
FT /tag= w
FT /label= PS19
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution I109V"
FT variation replace (2036, C)
FT /tag= x
FT /label= PS20
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (2039, T)
FT /tag= y
FT /label= PS21
FT /note= "Known single nucleotide polymorphism (SNP)"
FT intron 2056..2605
FT /tag= z
FT /number= 2
FT /cons_splice= (5'site:NO, 3'site:YES)
FT variation replace (2062, G)
FT /tag= aa
FT /label= PS22
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (2067, G)
FT /tag= ab
FT /label= PS23
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (2118, T)
FT /tag= ac
FT /label= PS24
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT variation replace (2170, A)
FT /tag= ad
FT /label= PS25
FT /note= "Known single nucleotide polymorphism (SNP)"
FT variation replace (2179, C)
FT /tag= ae
FT /label= PS26
FT /note= "Novel single nucleotide polymorphism (SNP)"
FT exon 2606..2758
FT /tag= af
FT /number= 3
FT variation replace (2611, A)
FT /tag= ag
FT /label= PS27
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution F120I"
FT variation replace (2635, C)
FT /tag= ah
FT /label= PS28
FT /note= "Novel single nucleotide polymorphism (SNP);
FT causes the amino acid substitution W128R"
FT variation replace (2659, A)
FT /tag= ai
FT /label= PS29
FT /note= "Novel single nucleotide polymorphism (SNP);
FT together with PS30 causes the amino acid
FT substitution V136I"
FT variation replace (2661, C)
FT /tag= aj
FT /label= PS30
```

FT /note= "Known single nucleotide polymorphism (SNP);  
 FT together with PS29 causes the amino acid  
 FT substitution V136I"  
 FT replace (2704, G)  
 FT /tag= ak  
 FT /label= PS31  
 FT /note= "Known single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution Q151E"  
 FT replace (2716, A)  
 FT /tag= al  
 FT /label= PS32  
 FT /note= "Novel single nucleotide polymorphism (SNP);  
 FT causes the amino acid substitution E155K"  
 FT 2759..2846  
 FT /tag= am  
 FT /number= 3  
 FT replace (2846, A)  
 FT /tag= an  
 FT /label= PS33  
 FT /note= "Known single nucleotide polymorphism (SNP)"  
 FT 2847..3007  
 FT /tag= ao  
 FT /number= 4  
 FT 3008..3440  
 FT /tag= ap  
 FT /number= 4  
 FT replace (3292, A)  
 FT variation

Query Match 99.0%; Score 40.6; DB 24; Length 6472;

Best Local Similarity 97.6%; Pred. No. 2.1e-06; Indels 0; Gaps 0;

Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGTGGGTGATTTCGTGTGTAATCGTGTCCCTG 41  
 |||||  
 Db 301 CTTTGTGGGTGATTTCGTGTGTAATCGTGTCCCTG 341

RESULT 4

AAD34213

ID AAD34213 standard; DNA; 9432 BP.

AC AAD34213;

XX 16-JUL-2002 (first entry)

XX Human cytochrome P450 2D6 (CYP2D6) gene.

XX Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
 KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
 KW gene; ds.  
 XX Homo sapiens.  
 OS WO200218638-A2.  
 PN 07-MAR-2002.  
 PD 27-AUG-2001; 2001WO-IB01544.  
 XX 30-AUG-2000; 2000GB-0021286.  
 PR (GEMI-) GEMINI GENOMICS PLC.  
 PA Risinger C, Andersson MK, Lewander T, Ollasson E;  
 XX WPI; 2002-329785/36.  
 DR

XX New sequence determination oligonucleotides, useful for detecting  
 PT polymorphic sites in a 5' flanking region of a CYP2D6 gene, as  
 PT hybridization probes, as components of diagnostic assays, or in  
 PT ligase-based sequence determination -  
 XX Example 3; Fig 1; 63pp; English.

PS

XX The invention relates to sequence determination oligonucleotides for  
 CC detecting polymorphic sites in a 5' flanking region of cytochrome P450  
 CC 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many  
 CC different xenobiotics. Human CYP2D6 gene is located on chromosome 22.  
 CC The oligonucleotides may be used as in situ hybridisation probes, in  
 CC ligase-based sequenced determination, as components of diagnostic assays,  
 CC as probes in sequence determination methods based on mismatches, as  
 CC hybridisation-based diagnostic assays, and as components of diagnostic  
 CC microarray. CYP2D6 is useful to predict variations in an individual's  
 CC ability to metabolise certain drugs. The present sequence is human  
 CC CYP2D6 gene.  
 XX

SQ Sequence 9432 BP; 1964 A; 2647 C; 2976 G; 1845 T; 0 other;

Query Match 99.0%; Score 40.6; DB 24; Length 9432;

Best Local Similarity 97.6%; Pred. No. 2.2e-06;

Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGTGGGTGATTTCGTGTGTAATCGTGTCCCTG 41  
 |||||  
 Db 920 CTTTGTGGGTGATTTCGTGTGTAATCGTGTCCCTG 960

RESULT 5

AAH26169

ID AAH26169 standard; DNA; 1669 BP.

XX AAH26169;

XX 17-SEP-2001 (first entry)

XX Human cytochrome P450 CYP2D6 gene promoter region.

XX Cytochrome P450; CYP2D6; promoter; drug metabolism; human;  
 KW diagnosis; therapy; ds.  
 XX Homo sapiens.

XX Key

FT primer\_bind Location/Qualifiers  
 complement (14..36)  
 /tag= a

FT primer\_bind /note= "amplification primer upf14"  
 337..358  
 /tag= b

FT primer\_bind /note= "sequence primer R1"  
 493..514  
 /tag= c

FT primer\_bind /note= "sequencing primer R2"  
 complement (565..577)  
 /tag= d

FT primer\_bind /note= "sequencing primer F2"  
 602..620  
 /tag= e

FT primer\_bind /note= "sequencing primer R3"  
 complement (968..988)  
 /tag= f

FT primer\_bind /note= "sequencing primer F3"  
 1124..1143  
 /tag= g

FT primer\_bind /note= "sequencing primer R4"  
 1605..1623  
 /tag= h

FT primer\_bind /note= "sequencing primer R5"  
 1650..1669  
 /tag= i

FT 5'UTR /note= "amplification primer upr1669"  
 1532..1619  
 /tag= j

FT CDS 1620..1669  
 /tag= k

FT /partial  
 /note= "5' region of CYP2D6 coding region"

```
XX PN WO200155432-A2.
XX XX
XX PD 02-AUG-2001.
XX PF 30-JAN-2001; 2001WO-EP00954.
XX PR 31-JAN-2000; 2000EP-0101889.
XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX PI Raimundo S, Zanger U;
XX DR WPI; 2001-457734/49.
XX PT A polynucleotide capable of hybridizing to CYP2D6 promoter useful for
XX PT the optimization of drug therapies using substrates of cytochrome P-450
XX PS
XX PS Claim 1; Fig 1; 41pp; English.
XX CC The present sequence is that of the promoter region of the human
XX CC cytochrome P450 CYP2D6 gene. The promoter region was amplified
XX CC by PCR from leucocyte DNA of over 50 individuals, and sequenced.
XX CC 8 Previously unknown single nucleotide polymorphisms (SNP) were
XX CC identified. These were at: base 36 (base -1584 according to the
XX CC Human Cytochrome P450 Allele Nomenclature), where the SNP was C to
XX CC G, occurring at an estimated frequency of approximately 20% in the
XX CC whole population, and resulting in increased enzyme activity;
XX CC position 194 (-1426), C to T, approximately 20% frequency, neutral
XX CC function; position 385 (-1235), A to G, approximately 50% frequency,
XX CC neutral function; position 620 (-1000), G to A, approximately 20%
XX CC frequency, neutral function; position 880 (-740), C to T,
XX CC approximately 30% frequency, unknown function; position 940 (-680),
XX CC G to A, approximately 30% frequency, unknown function; 1255 (-365),
XX CC G to A, rare, unknown function; and 1298 (-322), T to C, rare,
XX CC unknown function. The C to G mutation at -1584 bp is strongly
XX CC associated with lower metabolic ratios, and a molecular variant
XX CC polynucleotide having G at this position is claimed (see AAH26179).
XX CC The invention provides a method of diagnosing a disorder related to
XX CC reduced or enhanced capacity for clearance of CYP2D6 substrates
XX CC (antiarrhythmic, beta-adrenergic receptor-antagonist, tricyclic
XX CC antidepressant, selective serotonin reuptake inhibitor, neuroleptic,
XX CC opiate, cytostatic or amphetamine), or susceptibility to such a
XX CC disorder, by determining the presence of a mutation in the CYP2D6
XX CC promoter. The strong association of the common C to G mutation at
XX CC -1584 bp with increased enzyme activity significantly improves the
XX CC correlation between genotype and phenotype in the CYP2D6 polymorphism.
XX CC Testing for the mutation will allow the identification of intermediate
XX CC metabolizers and therefore allow quantitative predictions to be made
XX CC on in vivo drug metabolism capacity, thus providing a very potent
XX CC tool for improving the therapy of diseases with drugs that are
XX CC targets of the CYP2D6 gene product.
XX SQ Sequence 1669 BP; 413 A; 376 C; 534 G; 338 T; 8 other;
XX
XX Query Match 98.0%; Score 40.2; DB 22; Length 1669;
XX Best Local Similarity 95.1%; Pred. No. 2.2e-06;
XX Matches 39; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CTTTGTGCGGTGATTTCTCGRTGTGTAATCGTGTCCTG 41
XX DB 920 CTTTGTGCGGTGATTTCTTCGTGTGTAATCGTGTCCTG 960
XX
XX RESULT 6
XX ID AAH26179
XX AC AAH26179 standard; DNA; 1669 BP.
XX AC AAH26179;
XX DT 17-SEP-2001 (first entry)
XX XX
```

```
DE XX Human cytochrome P450 CYP2D6 gene promoter (G mutation at -1584 bp).
KW XX Cytochrome P450; CYP2D6; promoter; drug metabolism; human;
KW diagnosis; therapy; single nucleotide polymorphism; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Variation replace(36,G)
XX FT /*tag= a
XX FT /*frequency= "20%"
XX FT 5'UTR 1532..1619
XX FT /*tag= b
XX FT CDS 1620..1669
XX FT /*tag= c
XX FT /partial
XX FT /note= "5' region of CYP2D6 coding region"
XX PN WO200155432-A2.
XX XX
XX PD 02-AUG-2001.
XX PF 30-JAN-2001; 2001WO-EP00954.
XX PR 31-JAN-2000; 2000EP-0101889.
XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX PI Raimundo S, Zanger U;
XX DR WPI; 2001-457734/49.
XX PT A polynucleotide capable of hybridizing to CYP2D6 promoter useful for
XX PT the optimization of drug therapies using substrates of cytochrome P-450
XX PS
XX PS Claim 1(a); Page -; 41pp; English.
XX CC The present sequence is that of the promoter region of the human
XX CC cytochrome P450 CYP2D6 gene, which includes G at position 36
XX CC (base -1584 according to the Human Cytochrome P450 Allele
XX CC Nomenclature). The presence of C at position -1584 bp is a marker
XX CC for low enzyme activity, whereas there is strong association of G
XX CC at position -1584 bp with increased enzyme activity. The C to G
XX CC single nucleotide polymorphism occurs in approximately 20% of the
XX CC population. The invention provides a method of diagnosing a
XX CC disorder related to reduced or enhanced capacity for clearance of
XX CC CYP2D6 substrates (antiarrhythmic, beta-adrenergic receptor
XX CC antagonist, tricyclic antidepressant, selective serotonin reuptake
XX CC inhibitor, neuroleptic, opiate, cytostatic or amphetamine), or
XX CC susceptibility to such a disorder, by determining the presence of
XX CC a mutation in the CYP2D6 promoter. The novel variant forms of the
XX CC CYP2D6 gene provided by the invention provide the potential for the
XX CC development of a pharmacodynamic profile of drugs for a given
XX CC patient. The finding and characterization of variations in the
XX CC CYP2D6 gene, and diagnostic tests for the discrimination of
XX CC different alleles in human individuals, provide a very potent tool
XX CC for improving the therapy of diseases with drugs that are targets
XX CC of the CYP2D6 gene production, and whose metabolism is therefore
XX CC dependent on CYP2D6 activity.
XX CC Note: The present sequence is not shown in the specification but is
XX CC derived from the CYP2D6 promoter sequence given in the Sequence
XX CC Listing (see AAH26169).
XX SQ Sequence 1669 BP; 413 A; 376 C; 535 G; 338 T; 7 other;
XX
XX Query Match 98.0%; Score 40.2; DB 22; Length 1669;
XX Best Local Similarity 95.1%; Pred. No. 2.2e-06;
XX Matches 39; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CTTTGTGCGGTGATTTCTCGRTGTGTAATCGTGTCCTG 41
XX DB 920 CTTTGTGCGGTGATTTCTTCGTGTGTAATCGTGTCCTG 960
XX
XX CTTTGTGCGGTGATTTCTTCGTGTGTAATCGTGTCCTG 41
XX CTTTGTGCGGTGATTTCTTCGTGTGTAATCGTGTCCTG 960
```



Db 4302 TTTGTGCTGGTGATTTTGTATGTGTAATCGTGTTTGTG 4341

```
RESULT 7
ABK39958
ID ABL32556 standard; DNA; 5884 BP.
XX
AC ABL32556;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #20 strand 1.
XX
KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW Cytostatic; AldH6; CYP11A; CYP11B; CYP3A3; DPYD; EPHA2; OCLN; TXNRD1;
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200202806-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP07470.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154757/20.
XX
PT New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT useful for detecting cytosine methylation state of genes associated
PT with pharmacogenomics and for therapy of diseases e.g. cancer
XX
PS Claim 1; SEQ ID No 39; 24pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence at
CC least 18 bases in length of a segment of the chemically pretreated DNA
CC of genes associated with pharmacogenomics according to one of the
CC sequences of the genes ALDH6 (NM_000893), CYP11A (NM_000781), CYP11B1
CC (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
CC (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
CC MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
CC NM_019899) and their complementary sequences, or a sequence (Sl) chosen
CC from 87 sequences and their complements. The chemical pretreatment
CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)
CC into uracils. Also included are an oligomer (II) in particular an
CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
CC each case at least one base sequence having a length of 9 nucleotides
CC which hybridises to or is identical to a chemically pretreated DNA of
CC genes associated with pharmacogenomics and their complements, arranged in
CC an array for analysing diseases associated with the methylation state
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC of the 87 sequences. The oligomers may also be used as PCR primers.
CC The set of 87 nucleic acids and their complements is useful for diagnosis
CC and therapy of solid tumours and cancer. The present sequence
CC represents one the 87 DNA sequences or its complement.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5884 BP; 1259 A; 92 C; 1514 G; 3019 T; 0 other;

Query Match 77.1%; Score 31.6; DB 24; Length 5884;
Best Local Similarity 85.0%; Pred. No. 0.0077;
Matches 34; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 TTTGTGCTGGTGATTTTCTGCTGTGTAATCGTGTCCTCG 41
|||||
Db 4302 TTTGTGCTGGTGATTTTGTATGTGTAATCGTGTTTGTG 4341

RESULT 9
ABK36040
ID ABL32556 standard; DNA; 5884 BP.
XX
AC ABL32556;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 529.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianemic; cytosine methylation; neurotropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritis; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation
XX
PS Claim 1; SEQ ID NO 529; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 5884 BP; 1259 A; 92 C; 1514 G; 3019 T; 0 other;

Query Match 77.1%; Score 31.6; DB 24; Length 5884;
Best Local Similarity 85.0%; Pred. No. 0.0077;
Matches 34; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 TTTGTGCTGGTGATTTTCTGCTGTGTAATCGTGTCCTCG 41
|||||
Db 4302 TTTGTGCTGGTGATTTTGTATGTGTAATCGTGTTTGTG 4341

RESULT 9
ABK36040
ID ABL32556 standard; DNA; 2299 BP.
XX
AC ABL36040;
XX
DT 08-MAY-2002 (first entry)
XX
```



KW Human; immune system disease; cytosine methylation; antiasthmatic;  
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;  
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;  
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;  
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;  
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;  
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200200928-A2.  
XX  
XX  
PD 03-JAN-2002.  
XX  
XX 02-JUL-2001; 2001WO-EP07537.  
XX  
XX 30-JUN-2000; 2000DE-1032529.  
PR 01-SEP-2000; 2000DE-1043826.  
XX  
XX (EPG-) EPIGENOMICS AG.  
PA  
XX Olek A, Piepenbrock C, Berlin K;  
PI  
XX WPI; 2002-130909/17.  
DR  
XX  
XX Nucleic acid comprising fragment of chemically modified gene, useful  
PT for diagnosis and treatment of diseases associated with abnormal  
PT cytosine methylation -  
PT  
XX  
XX Claim 1; SEQ ID NO 530; 32pp + Sequence Listing; German.  
PS  
XX  
XX The present invention provides a number of human immune system associated  
CC genes which are modified by the methylation of cytosines. The sequences  
CC can be used in the diagnosis and treatment of immune system disorders,  
CC including eye diseases such as retinopathy, neovascular glaucoma and  
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid  
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,  
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel  
CC diseases. The present sequence is a gene of the invention.  
XX  
SQ Sequence 5884 BP; 1730 A; 92 C; 1381 G; 2681 T; 0 other;  
Query Match 57.6%; Score 23.6; DB 24; Length 5884;  
Best Local Similarity 72.5%; Pred. No. 12;  
Matches 29; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 1 CTTTGTGGTGATTTTCTGCGTGTGTAATCGTGCCCT 40  
||||| | | ||||| | : | ||||| |||||  
Db 1584 CTTTATAATAAATTTTCTACATATATAATCGTATCCCT 1545  
RESULT 12  
ABAI7450/c  
ID ABAI7450 standard; DNA; 6281 BP.  
XX  
AC ABAI7450;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Human nervous system related polynucleotide SEQ ID NO 9781.  
XX  
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulneryary;  
KW antiparkinsonian; antiskinking; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200159063-A2.



PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216680.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225271.  
PR 14-AUG-2000; 2000US-0225272.  
PR 14-AUG-2000; 2000US-0225273.  
PR 14-AUG-2000; 2000US-0225274.  
PR 14-AUG-2000; 2000US-0225275.  
PR 14-AUG-2000; 2000US-0225276.  
PR 14-AUG-2000; 2000US-0225277.  
PR 14-AUG-2000; 2000US-0225278.  
PR 14-AUG-2000; 2000US-0225279.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229349.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 06-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0232402.  
PR 14-SEP-2000; 2000US-0232403.  
PR 14-SEP-2000; 2000US-0232404.  
PR 14-SEP-2000; 2000US-0232405.  
PR 21-SEP-2000; 2000US-0232423.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-541565/60.  
Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
useful for preventing, diagnosing and/or treating nervous system  
cancers and metastases -  
Disclosure; SEQ ID NO 9780; 1701pp + Sequence Listing; English.  
The invention relates to novel genes (ABA11004-ABA21534) and proteins  
(AB14678-AB18001) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast



CC toxicity markers in drug screening and toxicity assays. The genes and  
CC gene expression information may be used as diagnostic markers for the  
CC prediction or identification of the physiological state of tissue or cell  
CC sample that has been exposed to a compound or agent. Hepatotoxicity  
CC is characterised by centrilobular necrosis and steatosis. The present  
CC sequence is an expressed sequence tag (EST) or cDNA derived from a gene  
CC which is differentially expressed in response to a hepatotoxic agent.

XX  
SQ Sequence 500 BP; 182 A; 115 C; 82 G; 121 T; 0 other;

Query Match 56.1%; Score 23; DB 24; Length 500;  
Best Local Similarity 70.7%; Pred No. 12;  
Matches 29; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 CTTTGTGCGGTGATTTTCTGCTGTGTGTAATCGTGTCCCTG 41  
|| ||||| ||| | |||| : |||| | || |||||  
Db 324 CTCGTGTGTGTCTGTCTGTGTGTGTCTTTTGTATGCCCTG 284

Search completed: February 11, 2003, 02:06:10  
Job time : 34.6596 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:51:09 ; Search time 118.283 Seconds  
(without alignments)  
5613.764 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_920\_960  
Perfect score: 41  
Sequence: 1 cttgtgtggtgatttctt.....crtgttaatcgtgtccctg 41

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum Match 0%  
Maximum DB seq length: 2000000000  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pln:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_othr:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.6	99.0	1180	14	BQ067695
2	25.4	62.0	516	12	BG657500
3	25.2	61.5	618	10	BE371772
4	25.2	61.5	1055	12	BF535613
5	25	61.0	499	13	BW114203
6	24.8	60.5	915	17	CNS04CLN

c 7	24.8	60.5	1824	12	BF128211
c 8	24.4	59.5	1036	12	BG787054
c 9	24.2	59.0	302	9	AA607596
c 10	24.2	59.0	308	14	F03457
c 11	24.2	59.0	353	14	N36563
c 12	24.2	59.0	387	9	AI052092
c 13	24.2	59.0	506	12	BF652228
c 14	24	58.5	383	10	BE623020
c 15	24	58.5	432	13	BI999135
c 16	24	58.5	879	17	CNS04011
c 17	23.8	58.0	908	13	BG918351
c 18	23.6	57.6	268	9	AA034947
c 19	23.6	57.6	285	17	BH274434
c 20	23.6	57.6	308	9	AA652245
c 21	23.6	57.6	311	9	AA483727
c 22	23.6	57.6	338	12	BF552720
c 23	23.6	57.6	340	17	AZ099911
c 24	23.6	57.6	347	9	A1758738
c 25	23.6	57.6	354	9	A1824445
c 26	23.6	57.6	361	14	R98298
c 27	23.6	57.6	361	17	R38696
c 28	23.6	57.6	373	14	R49180
c 29	23.6	57.6	376	14	BW761459
c 30	23.6	57.6	384	9	AI580308
c 31	23.6	57.6	412	9	AI888012
c 32	23.6	57.6	422	9	AI662312
c 33	23.6	57.6	464	17	BH050434
c 34	23.6	57.6	472	10	BE046391
c 35	23.6	57.6	475	17	BH063439
c 36	23.6	57.6	479	14	BH842494
c 37	23.6	57.6	496	17	AZ426410
c 38	23.6	57.6	509	9	AI590982
c 39	23.6	57.6	520	12	BE855424
c 40	23.6	57.6	555	14	BQ554798
c 41	23.6	57.6	564	12	BE857133
c 42	23.6	57.6	568	17	AZ342501
c 43	23.6	57.6	575	17	FR0022843
c 44	23.6	57.6	591	17	FR0043907
c 45	23.6	57.6	592	17	AZ051347

## ALIGNMENTS

RESULT 1  
BQ067695  
LOCUS BQ067695 1180 bp mRNA linear EST 02-APR-2002  
DEFINITION AGENCOURT\_6643098 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5768406  
5', mRNA sequence.  
ACCESSION BQ067695  
VERSION BQ067695.1 GI:19896741  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1180)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLAM12828 row: d column: 07  
High quality sequence start: 3  
High quality sequence stop: 536.  
Location/Qualifiers

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NIH-MGC <http://mgc.nci.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLAM12828 row: d column: 07  
High quality sequence start: 3  
High quality sequence stop: 536.  
Location/Qualifiers

```

source
1. .1180
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5768406"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH_MGC Library."
BASE COUNT      254 a   371 c   347 g   208 t
ORIGIN
Query Match      99.0%; Score 40.6; DB 14; Length 1180;
Best Local Similarity 97.6%; Pred. No. 0.00015;
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTTTGTGGGTGATTTCTGCTGCTGTAATCGTGCCCTG 41
|||||
Db 403 CTTTGTGGGTGATTTCTGCTGCTGTAATCGTGCCCTG 443
|||||

RESULT 2
BG657500
LOCUS              516 bp      mRNA      linear      EST 11-MAY-2001
DEFINITION        BG657500 TgVEG118 Tachyzoite cDNA Library Toxoplasma gondii
cDNA clone TgESTya20g06.y1 5', mRNA sequence.
ACCESSION         BG657500
VERSION           BG657500.1 GI:13797767
KEYWORDS          EST.
SOURCE            Toxoplasma gondii.
ORGANISM          Toxoplasma gondii.
REFERENCE         1 (bases 1 to 516)
AUTHORS           Tang,K., Cole,R., Fogarty,S., Sibley,L.D., Ajlola,J.A., White,M.,
Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M.,
Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter
,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy
,S., Maguire,L., Waterston,R. and Wilson,R.
TITLE             Toxoplasma EST Project
JOURNAL           Unpublished (2001)
COMMENT           Contact: Clifton, S.
                  Washington University School of Medicine
                  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                  Tel: 314 286 1800
                  Fax: 314 286 1810
                  Email: toxo@watson.wustl.edu
                  Contact David Sibley (toxoe@borcim.wustl.edu) for further
                  information relating to organism, libraries, or clone availability.
                  Seq primer: -40RP from Gibco
                  High quality sequence stop: 421.
                  Location/Qualifiers
                    1. .516
                      /organism="Toxoplasma gondii"
                      /strain="VEG"
                      /db_xref="taxon:5811"
                      /clone="TgSrya20g06.y1"
                      /clone_lib="TgVEG118 Tachyzoite cDNA Library"
                      /dev_stage="Tachyzoite"
                      /lab_host="DH10B"
                      /Note="Vector: pBluescript SK; Site_1: EcoRI; Site_2: XhoI
                      : This library was constructed by Keliang Tang, Robert
                      Cole, and L. David Sibley at Washington University. cDNAs
                      were synthesized from poly(A)+ RNA by oligod(T) priming,
                      size-selected and directionally cloned into the Uni-ZAP XR

```

```

lambda vector (Stratagene). The primary library was mass
excised as phagemids and rescued in SOLR cells. The
plasmid library was recovered from the SOLR cells and
transformed in mass into DH10B (GeneHog, Research Genetics
, Inc.) for sequencing. WARNING: This library may contain
a small percentage contaminants from human fibroblast
cells."
BASE COUNT      88 a   139 c   141 g   148 t
ORIGIN
Query Match      62.0%; Score 25.4; DB 12; Length 516;
Best Local Similarity 78.4%; Pred. No. 44;
Matches 29; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GGTGGGTGATTTCTGCTGCTGTAATCGTGCCCTG 41
|||||
Db 234 GGTGGGTGATTTCTGCTGCTGTAATCGTGCCCTG 270
|||||

RESULT 3
BE371772/c
LOCUS              618 bp      mRNA      linear      EST 21-JUL-2000
DEFINITION        BE371772 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3587125 5',
mRNA sequence.
ACCESSION         BE371772
VERSION           BE371772.1 GI:9317135
KEYWORDS          EST.
SOURCE            house mouse.
ORGANISM          Mus musculus
REFERENCE         1 (bases 1 to 618)
AUTHORS           NIH-MGC http://mac.nci.nih.gov/
TITLE             National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL           Unpublished (1999)
COMMENT           Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-r@mail.nih.gov
                  Tissue procurement: Gilbert Smith, Ph.D.
                  cDNA Library Preparation: Life Technologies, Inc.
                  cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: LLAM8749 row: m column: 14
                  High quality sequence stop: 614.
                  Location/Qualifiers
                    1. .618
                      /organism="Mus musculus"
                      /strain="CZECH II"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:3587125"
                      /clone_lib="NCI_CGAP_Lu29"
                      /tissue_type="spontaneous tumor, metastatic to mammary.
                      Stem cell origin."
                      /lab_host="DH10B"
                      /Note="organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally. Primer: oligo dT.
                      Library constructed by Life Technologies. Investigator
                      providing samples: Gilbert Smith, NIH"
BASE COUNT      181 a   169 c   151 g   117 t
ORIGIN
Query Match      61.5%; Score 25.2; DB 10; Length 618;
Best Local Similarity 75.0%; Pred. No. 54;
Matches 30; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 2 TTTGTGGGTGATTTCTGCTGCTGTAATCGTGCCCTG 41
|||||
Db 563 TTTGTGTCTCCCTTGTGCGGTGTAATGTAATGTCGCTG 524
|||||

RESULT 4

```

Seq primer: -21M13 Forward	
High quality sequence stop: 499	
POLYA=Yes.	
FEATURES	
source	Location/Qualifiers
	1. .499
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="niaEST:L0801B08-3"
	/db_xref="taxon:10090"
	/clone="L0801B08"
	/clone_lib="NIA Mouse Newborn Brain cDNA Library"
	/tissue_type="Newborn Brain"
	/dev_stage="Newborn"
	/lab_host="Dh10B"
	/note="Organ: brain; Vector: pSPORT1 (Invitrogen); Site:1: SalI; Site:2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research program, NIH (http://lgsum.grc.nia.nih.gov/cDNA). Double-stranded cDNAs were synthesized with an Oligo(dT) primer [Invitrogen: 5'-pGATGTTCTAGATCGGCGCGCCCTTTTTTTTTTTT-3'] from 48 microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal3 (Ref. Development 127:1737-1749 (2000) [PMID:10725249]), purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were digested with SalI and NotI enzymes, and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. Coli host was transformed with ligation mixture by the standard chemical method. The average insert size is about 1.9 kb. The library was constructed by Yulan Piao (NIA)."
BASE COUNT	156 a 122 c 215 g 6 t
ORIGIN	
Query Match	61.0%; Score 25; DB 13; Length 499;
Best Local Similarity	75.6%; Pred. No. 61;
Matches	31; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY	1 CTTTGTGGGTGATTTCCTGCRGTGTAATCGTGTCCTCG 41 
Db	289 CTTTCGGGGGCTCTTTCCTGCTGGGTTCTCTGCTCTG 249
RESULT 6	
CNS04CLN/c	
LOCUS	CNS04CLN 915 bp DNA linear GSS 21-MAY-2000
DEFINITION	Tetraodon nigroviridis genome survey sequence PUC-ori and of clone 099D06 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION	AL284612.1 GI:8023003
VERSION	GSS: genome survey sequence.
KEYWORDS	Tetraodon nigroviridis.
SOURCE	Tetraodon nigroviridis.
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE	1 (bases 1 to 915)
AUTHORS	Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 915)
AUTHORS	Roest-Crolius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Unpublished

```

REFERENCE 3 (bases 1 to 915)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

FEATURES
source
1..915
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="099P06"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG099DB03SP1-end :
PUC-ori"
BASE COUNT 250 a 242 c 277 g 133 t 13 others
ORIGIN
Query Match 60.5%; Score 24.8; DB 17; Length 915;
Best Local Similarity 72.2%; Pred. No. 82;
Matches 26; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 CTTTGTGCGTGATTTCTTCGRTGTGTAATCGTGT 36
||||:|||| ||| |:|||||:||||| : ||||
Db 280 CTTTGTGCTGTGKTCTCGCGTGTGKTGTGT 245

RESULT 7
BF128211/c
LOCUS BF128211 1824 bp mRNA linear EST 24-OCT-2000
DEFINITION 601810436R1 NTH_MGC_46 Homo sapiens cDNA clone IMAGE:4053536 3',
mRNA sequence.
ACCESSION BF128211.1 GI:10967251
VERSION BF128211.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1824)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC892 row: g column: 09
High quality sequence start: 7
High quality sequence stop: 31.
Location/Qualifiers
1..1824
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4053536"
/clone_lib="NTH_MGC_46"
/tissue_type="leiomyosarcoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pOT87; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

```

BASE COUNT 700 a 472 c 402 g 249 t 1 others
ORIGIN
Query Match 60.5%; Score 24.8; DB 12; Length 1824;
Best Local Similarity 76.3%; Pred. No. 95;
Matches 29; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 TTTGTGCGGTGATTTCTTCGRTGTGTAATCGTGCC 39
|||||:||||| ||| |:||||| :||||| |
Db 1336 TTTGTGCGGTGATTTCTTCGRTGTGTAATCGTGACAC 1299

RESULT 8
BG787054/c
LOCUS BG787054 1036 bp mRNA linear EST 20-MAY-2001
DEFINITION SEAMC007011 Sea urchin primary mesenchyme cell cDNA library
Strongylocentrotus purpuratus cDNA clone PM990802-02-0457 5', mRNA
sequence.
ACCESSION BG787054.1 GI:14158067
VERSION BG787054.1
KEYWORDS EST.
SOURCE Strongylocentrotus purpuratus.
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
REFERENCE 1 (bases 1 to 1036)
AUTHORS Zhu,X., Mahairas,G., Illies,M.R., Cameron,R.A., Davidson,E.H. and
Ettensohn,C.A.
TITLE A large scale analysis of mRNAs expressed by primary mesenchyme
cells of the sea urchin embryo
JOURNAL Development 128 (13), 2615-2627 (2001)
MEDLINE 21384984
COMMENT Contact: Ettensohn CA
Dept. Biol. Sci.
Carnegie Mellon University
4400 Fifth Avenue, Pittsburgh, PA 15213, USA
Tel.: +1 412 268 5849
Email: ettensohn@andrew.cmu.edu.
Location/Qualifiers
1..1036
/organism="Strongylocentrotus purpuratus"
/db_xref="taxon:7668"
/clone="PM990802-02-0457"
/clone_lib="Sea urchin primary mesenchyme cell cDNA
library"
/tissue_type="embryo"
/cell_type="primary mesenchyme cells"
/lab_host="E.coli"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; oligo
dT priming from poly A+ RNA, directionally cloned"
BASE COUNT 349 a 223 c 197 g 263 t 4 others
ORIGIN
Query Match 59.5%; Score 24.4; DB 12; Length 1036;
Best Local Similarity 77.8%; Pred. No. 1.2e+02;
Matches 28; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TTTGTGCGGTGATTTCTTCGRTGTGTAATCGTGTC 37
|||||:||||| ||| |:||||| :||||| |
Db 409 TTTGTGCGGTGATTTCTTCGRTGTGTAATCGTTTC 374

RESULT 9
AA607596
LOCUS AA607596 302 bp mRNA linear EST 30-SEP-1997
DEFINITION VO07h12.rl Stratagene mouse skin (#937313) Mus musculus cDNA clone
IMAGE:1040807 5', mRNA sequence.
ACCESSION AA607596
VERSION AA607596.1 GI:2455031
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus. 1 (bases 1 to 302) Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterson,R.	
TITLE	The WashU-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Marra M/Mouse EST Project  WashU-HHMI Mouse EST Project  Washington University School of MedicineP 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810  Email: mouseest@watson.wustl.edu  This clone is available royalty-free through LLNL ; contact the IMAGE Consortium ( <a href="#">info@image.llnl.gov</a> ) for further information. MGID:360831
Seq primer:	-28ml3 rev1 ET from Amershams
High quality sequence stop:	213.
Location/Qualifiers	
Source	1..302 /organism="Mus musculus" /strain="C57BL/6" /db_xref="taxon:10090" /cloned_image=I040807" /clone_lib="Stratagene mouse skin (#937313)" /sex="females" /tissue_type="whole skin" /dev_stage="11 weeks old" /lab_host="SOLR (kanamycin resistant)" /note="Organ: skin; Vector: pBluescript SK-; Site.1: EcoRI ; Site.2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Whole skin from 11 week old C57BL/6 female mice. Average insert size: 1.0 kb; Uni-ZAP XR vector; ~5' adaptor sequence: 5' GAATTCGCACAG 3' -3' adaptor sequence: 5' CTCGAGTGTGGTTTTTTTTTTTTT 3'
BASE COUNT	76 a   61 c   70 g   95 t
ORIGIN	
Query Match	59.0%; Score 24.2; DB 9; Length 302;
Best Local Similarity	74.4%; Pred. No. 1.le+02;
Matches 29; Conservative 1; Mismatches 9; Indels 0; Gaps 0;	
QY	2 TTTCTGTGGTGATTTCCTCGTGTGAATCGTCCCT 40                 :             Db 111 TGCTGTGTGTGTGTGTGTGTGTGTGAATGTGTCCCT 149
RESULT 10	
F03457/c	
LOCUS	F03457 308 bp mRNA linear EST l8-FEB-1995
DEFINITION	HSC1YA062 normalized infant brain cdna Homo sapiens cDNA clone
C-l ya06 3 ', mRNA sequence.	
ACCESSION	F03457
VERSION	F03457.1 GI:668831
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 308) Auffray,C., Behar,O., Bois,F., Bouchier,C., da Silva,C., Devignes .M.D., Duprat,S., Houllatte,R., Jumeau.M.N., Lamy.B., Lorenzo.F., Mitchell.H., Mariage-Samon.R., Pietu.G., Pouillot.Y., Sebastiani-Kabatchis.C. and Tessier,A. IMAGE: molecular integration of the analysis of the human genome and its expression C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995) 95277534 Contact: Genethon	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
CONTACT	



Durham, NC 27708-1000  
Tel: 919 613 8159  
Fax: 919 613 8177  
Email: chauser@duke.edu.

Location/Qualifiers  
1. .432  
/organism="Chlamydomonas reinhardtii"  
/strain="CC-1690 wild type mt+ 21gr"  
/db\_xref="taxon:3055"  
/clone\_lib="C. reinhardtii CC-1690, Stress II (normalized  
, lambda zap II")  
/note="vector: pBluescript II SK-; Site\_1: EcoRI; Site\_2:  
XhoI; Stress condition II library, constructed by John  
Davies and Jeffrey McDermott, combines cDNAs from CC-1690  
cells grown to mid-log phase in TAP (NH4 - containing)  
and shifted to TAP - NO3- (24hrs); H2 production  
conditions (0, 12hr, 24hr) see Mellis et al., (2000) Plant  
Phys. 122: 127-135; TAP + H2O2 (1, 12, 24 hr); TAP +  
sorbitol (1, 2, 6, 24 hr); TAP + Cd (1, 2, 6, 24 hr).  
polyA mRNA was purified from each sample, pooled and cDNA  
synthesized. The cDNA was directionally cloned into lambda  
zap II (Stratagene) in the EcoRI (5') and XhoRI (3')  
sites. pBluescript II SK- plasmids were excised from the  
lambda ZAP clones by superinfection with ExAssist  
(Stratagene) phage. The library was normalized using  
method 4 described in Bonaldo et al., (1996) Genome  
Research 6: 791-806."

```

BASE COUNT      114 a      131 c      99 g      88 t
ORIGIN
Query Match      58.5%; Score 24; DB 13; Length 432;
Best Local Similarity 79.4%; Pred. No. 1.4e+02;
Matches 27; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY      5 GTGTGGGTGATTTCCTGCTGTGTAACGTGTC 38
          | | | | | | | | | | | | | | | |
Db      275 GTGTGGGTGCGCTGTGTGTGAACGTGACC 242

Search completed: February 11, 2003, 03:31:24
Job time : 123.283 secs

```





Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	21.6	52.7	50000	4	US-09-146-053-4	Sequence 4, Appli
2	21.4	52.2	392	1	US-08-253-155A-60	Sequence 50, Appli
3	21.2	51.7	2169	3	US-08-806-326-5	Sequence 5, Appli
4	21	51.2	10684	3	US-08-618-100B-3	Sequence 3, Appli
5	20.8	50.7	1851	2	US-08-414-657D-20	Sequence 20, Appl
6	20.8	50.7	1851	4	US-09-135-080-5	Sequence 5, Appli
7	20.6	50.2	70000	4	US-09-851-896-3	Sequence 3, Appli
8	20.4	49.8	57	1	US-08-222-177A-62	Sequence 62, Appl
9	20.4	49.8	262	1	US-08-222-177A-46	Sequence 46, Appl
10	20.4	49.8	564	1	US-08-117-362-32	Sequence 32, Appl
11	20.4	49.8	564	1	US-08-486-924-32	Sequence 32, Appl
12	20.4	49.8	900	4	US-09-641-638-439	Sequence 439, Appl
13	20.4	49.8	1050	1	US-08-599-252-81	Sequence 81, Appl
14	20.4	49.8	1050	1	US-08-436-074-54	Sequence 54, Appl
15	20.4	49.8	1050	5	PC7-US96-06352-81	Sequence 81, Appl
16	20.4	49.8	1050	5	PC7-US96-06583-81	Sequence 81, Appl
17	20.4	49.8	2875	3	US-08-458-434A-4	Sequence 4, Appli
18	20.4	49.8	3796	2	US-08-762-308-11	Sequence 11, Appl
19	20.4	49.8	3796	4	US-09-844-634-10	Sequence 10, Appl
20	20.4	49.8	3813	2	US-08-650-000-3	Sequence 3, Appli
21	20.4	49.8	3813	6	5395760-3	Patent No. 5395760
22	20.4	49.8	4718	3	US-08-936-135-9	Sequence 9, Appli
23	20.4	49.8	4733	3	US-08-936-135-11	Sequence 11, Appl
24	20.4	49.8	4769	3	US-08-936-135-13	Sequence 13, Appl
25	20.4	49.8	4771	2	US-08-866-650-2	Sequence 2, Appli
26	20.4	49.8	4771	2	US-09-021-287-2	Sequence 2, Appli
27	20.4	49.8	4771	4	US-09-240-473-2	Sequence 2, Appli



Query Match 51.2%; Score 21; DB 3; Length 10684;  
Best Local Similarity 69.2%; Pred. No. 25;  
Matches 27; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2 TTGTGTTGGTGAATTTCTGCGTGTGTAATCGTGCCT 40  
| | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 3733 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCT 3771

RESULT 5

US-08-414-657D-20  
; Sequence 20, Application US/08414657D  
; Patent No. 5861283

GENERAL INFORMATION:  
; APPLICANT: Levitt, Pat  
; APPLICANT: Pimenta, Aurea  
; APPLICANT: Fischer, Itzhak  
; APPLICANT: Zhukareva, Victoria  
; TITLE OF INVENTION: Lmbic System-Associated Membrane  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 997 Lenox Drive, Building 3, Suite 210  
; CITY: Lawrenceville  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08543

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/414,657D  
; FILING DATE: 31-MAR-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bloom, Allen  
; REGISTRATION NUMBER: 29,135  
; REFERENCE/DOCKET NUMBER: 317743-102  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 609-520-3214  
; TELEFAX: 609-520-3259  
; TELEX:

INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1851 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 877..1032  
; OTHER INFORMATION:  
US-08-414-657D-20

Query Match 50.7%; Score 20.8; DB 2; Length 1851;  
Best Local Similarity 73.5%; Pred. No. 21;  
Matches 25; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 TTGTGTTGGTGAATTTCTGCGTGTGTAATCGTG 35  
| | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1078 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAGTG 1111

RESULT 6

US-09-135-080-5  
; Sequence 5, Application US/09135080

Patent No. 6423827  
; GENERAL INFORMATION:  
; APPLICANT: Levitt, Pat R.  
; APPLICANT: Pimenta, Aurea  
; APPLICANT: Fischer, Itzhak  
; APPLICANT: Zhukareva, Victoria  
; TITLE OF INVENTION: Lmbic System-Associated Membrane  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dechert Price & Rhoads  
; STREET: 997 Lenox Drive, Building 3, Suite 210  
; CITY: Lawrenceville  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08543

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/135,080  
; FILING DATE: 17-AUG-1998  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/414,657  
; FILING DATE: 31-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bloom, Allen  
; REGISTRATION NUMBER: 29,135  
; REFERENCE/DOCKET NUMBER: 317743-102A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 609-620-3214  
; TELEFAX: 609-620-3259  
; TELEX:

INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1851 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 877..1032  
; OTHER INFORMATION:  
US-09-135-080-5

Query Match 50.7%; Score 20.8; DB 4; Length 1851;  
Best Local Similarity 73.5%; Pred. No. 21;  
Matches 25; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 TTGTGTTGGTGAATTTCTGCGTGTGTAATCGTG 35  
| | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 1078 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAGTG 1111

RESULT 7

US-09-851-896-3/c  
; Sequence 3, Application US/09851896  
; Patent No. 6410325

GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Susan M. Freier  
; APPLICANT: Andrew T. Watt  
; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEP  
; FILE REFERENCE: RTS-0220  
; CURRENT APPLICATION NUMBER: US/09/851,896  
; CURRENT FILING DATE: 2001-05-08  
; NUMBER OF SEQ ID NOS: 89  
; SEQ ID NO 3  
; LENGTH: 70000



```
;
; OTHER INFORMATION: /standard_name= "Only one strand sequenced"
; PUBLICATION INFORMATION:
; AUTHORS: Weber, J. L.
; AUTHORS: Kwitek, A. E.
; AUTHORS: May, P. E.
; TITLE: Dinucleotide repeat polymorphism at the
; TITLE: D12S43 locus
; JOURNAL: Nucleic Acids Res.
; VOLUME: 18
; PAGES: 4637-
; DATE: 1990
; PUBLICATION INFORMATION:
; AUTHORS: Weber, James L.
; AUTHORS: May, Paula E.
; TITLE: Abundant Class of Human DNA Polymorphisms
; TITLE: Which Can Be Typed Using the Polymerase Chain
; TITLE: Reaction
; JOURNAL: Am. J. Hum. Genet.
; VOLUME: 44
; PAGES: 388-396
; DATE: 1989
; US-08-222-177A-46

Query Match          49.8%; Score 20.4; DB 1; Length 262;
Best Local Similarity 67.5%; Pred. No. 20;
Matches 27; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 TTGTGTCGTGATTTCTGCRGTGTAATCGTGTCCTG 41
   | ||||| ||| | | | : ||||| ||||| |||
Db 200 TGTGTGTGTGTGTGTGTGTGTGTCGCCGGTATCTG 161

RESULT 10
US-08-117-362-32/c
; Sequence 32, Application US/08117362
; Patent No. 5595872
; GENERAL INFORMATION:
; APPLICANT: Wetterau II, John R.
; APPLICANT: Sharp, Daru Y.
; APPLICANT: Gregg, Richard E.
; TITLE OF INVENTION: MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burton Rodney
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08117362
; FILING DATE: 03-SEP-1993
; PRIORITY/AGENT INFORMATION:
; NAME: Gaul, Timothy J.
; REGISTRATION NUMBER: 33,111
; REFERENCE/DOCKET NUMBER: DC21b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-5901
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 286..347
; US-08-117-362-32/c

Query Match          49.8%; Score 20.4; DB 1; Length 564;
Best Local Similarity 67.5%; Pred. No. 24;
Matches 27; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 TTGTGTCGTGATTTCTGCRGTGTAATCGTGTCCTG 41
   | ||||| ||| | | | : ||||| ||||| |||
Db 339 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCG 300

RESULT 12
US-09-641-638-439/c
; Sequence 439, Application US/09641638
```

```
;
; NAME/KEY: repeat_region
; LOCATION: 286..347
; US-08-117-362-32

Query Match          49.8%; Score 20.4; DB 1; Length 564;
Best Local Similarity 67.5%; Pred. No. 24;
Matches 27; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 TTGTGTCGTGATTTCTGCRGTGTAATCGTGTCCTG 41
   | ||||| ||| | | | : ||||| ||||| |||
Db 339 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCG 300

RESULT 11
US-08-486-924-32/c
; Sequence 32, Application US/08486924
; Patent No. 5789197
; GENERAL INFORMATION:
; APPLICANT: Wetterau II, John R.
; APPLICANT: Sharp, Daru Y.
; APPLICANT: Gregg, Richard E.
; TITLE OF INVENTION: MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burton Rodney
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: U.S.A.
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08486,924
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,362
; FILING DATE: 03-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaul, Timothy J.
; REGISTRATION NUMBER: 33,111
; REFERENCE/DOCKET NUMBER: DC21b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-5901
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: repeat_region
; LOCATION: 286..347
; US-08-486-924-32

Query Match          49.8%; Score 20.4; DB 1; Length 564;
Best Local Similarity 67.5%; Pred. No. 24;
Matches 27; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 2 TTGTGTCGTGATTTCTGCRGTGTAATCGTGTCCTG 41
   | ||||| ||| | | | : ||||| ||||| |||
Db 339 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCG 300

RESULT 12
US-09-641-638-439/c
; Sequence 439, Application US/09641638
```

```
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bouqueloret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annie
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641.638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502.330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133.200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275.267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119.917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 439
; LENGTH: 900
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 12-776-259 : polymorphic base A or G
; NAME/KEY: misc.binding
; LOCATION: 481..500
; OTHER INFORMATION: 12-776-259.misl, potential
; NAME/KEY: misc.binding
; LOCATION: 502..521
; OTHER INFORMATION: 12-776-259.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 243..263
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 674..692
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc.binding
; LOCATION: 489..513
; OTHER INFORMATION: 12-776-259 potential probe
; US-09-641-638-439
Query Match 49.8%; Score 20.4; DB 4; Length 900;
Best Local Similarity 67.5%; Pred. No. 26;
Matches 27; Conservative 1; Mismatches 12; Indels 0; Gaps 0;
QY 2 TTGTGTGGGTGATTTCTGTCGTGTAATCGTGTCCTG 41
| | | | | | | | | | | | | | | | | | | | |
Db 172 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 133
RESULT 13
US-08-599-252-81
; Sequence 81, Application US/08599252
; Patent No. 5705343
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIRKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
```

```
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,252
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 81:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1050 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-599-252-81
Query Match 49.8%; Score 20.4; DB 1; Length 1050;
Best Local Similarity 67.5%; Pred. No. 27;
Matches 27; Conservative 1; Mismatches 12; Indels 0; Gaps 0;
QY 2 TTGTGTGGGTGATTTCTGTCGTGTAATCGTGTCCTG 41
| | | | | | | | | | | | | | | | | | | | |
Db 521 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 560
RESULT 14
US-08-436-074-54
; Sequence 54, Application US/08436074
; Patent No. 5753438
; GENERAL INFORMATION:
; APPLICANT: DRAYNA, DENNIS T.
; APPLICANT: FEDER, JOHN N.
; APPLICANT: GNIRKE, ANDREAS
; APPLICANT: KIMMEL, BRUCE E.
; APPLICANT: THOMAS, WINSTON J.
; APPLICANT: WOLFF, ROGER K.
; TITLE OF INVENTION: METHOD TO DIAGNOSE HEREDITARY
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/436,074
; FILING DATE: 08-MAY-1995
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 9053-0001.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
```







GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2003, 03:31:30 ; Search time 3.72137 Seconds  
(without alignments)  
5222.300 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_920\_960  
Perfect score: 41  
Sequence: 1 ctttgtggtgatttct.....crttgttaatcgtgtccctg 41

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 408267 seqs, 237001491 residues

Total number of hits satisfying chosen parameters: 816534

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US07\_PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	DB ID	Description	
c	1	23	56.1	500	10	US-09-917-800A-26
	2	23	53.7	366	9	US-10-079-623-7
	3	22	53.7	5629	10	US-09-764-847-1162
	4	22	53.7	6265	10	US-09-764-847-1161
	5	22	53.7	203654	10	US-09-820-905-3
	6	21.8	53.2	5591	10	US-09-764-846-309
	7	21.6	52.7	539	10	US-09-864-761-12977
c	8	21.6	52.7	116592	10	US-09-818-512-3
	9	21.4	52.2	484	10	US-09-864-761-1146
	10	21.4	52.2	8894	10	US-09-764-847-1606
	11	21.4	52.2	22118	9	US-09-799-462A-16
c	12	21.4	52.2	22118	9	US-10-125-767-16
	13	21.4	52.2	22118	9	US-09-815-981-5
	14	21.4	52.2	659158	9	US-09-771-208-20
	15	21.2	51.7	73308	10	US-09-954-456-2276
c	16	21	51.2	243	10	US-09-880-107-3478
	17	21	51.2	361	9	US-09-796-692-3950
	18	21	51.2	398	10	US-09-867-701-7571
	19	21	51.2	528	9	US-09-796-692-3999

c 20	21	51.2	1686	9	US-10-114-893-37	Sequence 37, Appl
c 21	21	51.2	1834	10	US-09-974-300-1836	Sequence 1836, Ap
c 22	21	51.2	2682	9	US-10-098-841-99	Sequence 99, Appl
c 23	21	51.2	14537	10	US-09-962-832-232	Sequence 232, Appl
c 24	21	51.2	368004	10	US-09-949-654-3	Sequence 3, Appl1
c 25	20.8	50.7	793	10	US-09-878-574-4304	Sequence 4304, Ap
c 26	20.8	50.7	2607	9	US-09-991-053-11	Sequence 11, Appl
c 27	20.8	50.7	55155	10	US-09-735-933-3	Sequence 3, Appl1
c 28	20.8	50.7	659158	9	US-09-771-208-20	Sequence 20, Appl
c 29	20.6	50.2	582	10	US-09-864-761-16144	Sequence 16144, A
c 30	20.6	50.2	886	10	US-09-867-550-1009	Sequence 1009, Ap
c 31	20.6	50.2	2787	10	US-09-764-864-412	Sequence 412, App
c 32	20.6	50.2	3257	10	US-09-764-847-1410	Sequence 1410, Ap
c 33	20.6	50.2	3257	10	US-09-764-847-1411	Sequence 1411, Ap
c 34	20.6	50.2	19969	9	US-10-190-593-3	Sequence 3, Appl1
c 35	20.4	49.8	281	9	US-10-046-935-1107	Sequence 1107, Ap
c 36	20.4	49.8	281	9	US-09-878-178-1107	Sequence 1107, Ap
c 37	20.4	49.8	306	10	US-09-983-965-1180	Sequence 1180, Ap
c 38	20.4	49.8	369	9	US-10-015-219-320	Sequence 320, App
c 39	20.4	49.8	369	10	US-09-777-564-320	Sequence 320, App
c 40	20.4	49.8	370	10	US-09-864-761-18793	Sequence 18793, A
c 41	20.4	49.8	390	10	US-09-864-761-3024	Sequence 3024, Ap
c 42	20.4	49.8	426	10	US-09-917-800A-279	Sequence 279, App
c 43	20.4	49.8	506	10	US-09-962-832-151	Sequence 151, App
c 44	20.4	49.8	584	10	US-09-864-761-9954	Sequence 9954, Ap
c 45	20.4	49.8	721	10	US-09-917-800A-881	Sequence 881, App

ALIGNMENTS

RESULT 1

US-09-917-800A-26/c

Sequence 26, Application US/09917800A

Patent No. US20020119462A1

GENERAL INFORMATION:

APPLICANT: Mendrick, Donna

APPLICANT: Porter, Mark

APPLICANT: Johnson, Kory

APPLICANT: Castile, Arthur

APPLICANT: Elashoff, Michael

APPLICANT: Gene Logic, Inc.

TITLE OF INVENTION: Molecular Toxicology Modeling

FILE REFERENCE: 44921-5038-US

CURRENT APPLICATION NUMBER: US/09/917,800A

CURRENT FILING DATE: 2001-07-31

PRIOR APPLICATION NUMBER: US 60/222,040

PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: US 60/222,880

PRIOR FILING DATE: 2000-11-02

PRIOR APPLICATION NUMBER: US 60/290,029

PRIOR FILING DATE: 2001-05-11

PRIOR APPLICATION NUMBER: US 60/290,645

PRIOR FILING DATE: 2001-05-15

PRIOR APPLICATION NUMBER: US 60/292,336

PRIOR FILING DATE: 2001-05-22

PRIOR APPLICATION NUMBER: US 60/295,798

PRIOR FILING DATE: 2001-06-06

PRIOR APPLICATION NUMBER: US 60/297,457

PRIOR FILING DATE: 2001-06-13

PRIOR APPLICATION NUMBER: US 60/298,884

PRIOR FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: US 60/303,459

PRIOR FILING DATE: 2001-07-09

NUMBER OF SEQ ID NOS: 1740

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 26

LENGTH: 500

TYPE: DNA

ORGANISM: Rattus norvegicus

FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA799812

US-09-917-800A-26

```
Query Match          56.1%; Score 23; DB 10; Length 500;
Best Local Similarity 70.7%; Pred. No. 7.9;
Matches 29; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 CTTGTGCGGTGATTTCTGCRGTGTGTAATCGTGCCCTG 41
   ||||| ||| ||||| ||||| ||||| ||||| |||||
Db 324 CTTGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTG 284

RESULT 2
US-10-079-623-7
; Sequence 7, Application US/10079623
; Patent No. US20020169302A1
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka J.
; APPLICANT: Glenn, Matthew
; APPLICANT: Grigor, Murray R.
; APPLICANT: Molenaar, Adrian J.
; TITLE OF INVENTION: Compositions isolated from bovine
; FILE REFERENCE: 11000.1044c3
; CURRENT APPLICATION NUMBER: US/10/079,623
; CURRENT FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Bovine
US-10-079-623-7

Query Match          53.7%; Score 22; DB 9; Length 366;
Best Local Similarity 70.0%; Pred. No. 17;
Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2 TTGTGTCGGTGTATTTCTGCRGTGTGTAATCGTGCCCTG 41
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 TTGTGTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 138

RESULT 3
US-09-764-847-1162
; Sequence 1162, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1162
; LENGTH: 5629
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1162

Query Match          53.7%; Score 22; DB 10; Length 5629;
Best Local Similarity 70.0%; Pred. No. 29;
Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2 TTGTGTCGGTGTATTTCTGCRGTGTGTAATCGTGCCCTG 41
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3904 TGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3943

RESULT 4
US-09-764-847-1161
; Sequence 1161, Application US/09764847
; Patent No. US20020132767A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1161
; LENGTH: 6265
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-847-1161

Query Match          53.7%; Score 22; DB 10; Length 6265;
Best Local Similarity 70.0%; Pred. No. 29;
Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 2 TTGTGTCGGTGTATTTCTGCRGTGTGTAATCGTGCCCTG 41
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3904 TGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3943

RESULT 5
US-09-820-905-3
; Sequence 3, Application US/09820905
; Patent No. US20020142938A1
; GENERAL INFORMATION:
; APPLICANT: Yan, Chunhua
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001199
; CURRENT APPLICATION NUMBER: US/09/820,905
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 203654
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)---(203654)
; OTHER INFORMATION: n = A,T,C or G
US-09-820-905-3

Query Match          53.7%; Score 22; DB 10; Length 203654;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 28; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 CTTGTGTCGGTGTATTTCTGCRGTGTGTAATCGTGCCCT 40
   ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 93447 CTTGTGTCGTGCTGCTGCTGCTGCTGCTGCTGCT 93486

RESULT 6
US-09-764-846-309
; Sequence 309, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT212
; CURRENT APPLICATION NUMBER: US/09/764,846
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 309
; LENGTH: 5591
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-764-846-309

Query Match          53.2%; Score 21.8; DB 10; Length 5591;
Best Local Similarity 74.3%; Pred. No. 34;
Matches 26; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 TTGTGTGGTGATTTCTGCRGTGTAATCGTGT 36
||||| ||| ||||| ||||| |||||
Db 996 TTGTGTGTGTGTTTGTGTGTGTGTGTGTGT 1030

RESULT 7
US-09-864-761-12977
; Sequence 12977, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 12977
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010634.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.59
```

```
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.56
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.61
US-09-864-761-12977

Query Match          52.7%; Score 21.6; DB 10; Length 539;
Best Local Similarity 71.1%; Pred. No. 26;
Matches 27; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 CTTTGTGGTGATTTCTGCRGTGTAATCGTGCC 38
||||| ||| ||||| ||||| |||||
Db 454 CTGTGTGTGTGTGTGTGTGTGTGTGTGTTC 491

RESULT 8
US-09-818-512-3
; Sequence 3, Application US/09818512
; Patent No. US20020142416A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001192
; CURRENT APPLICATION NUMBER: US/09/818,512
; CURRENT FILING DATE: 2001-03-28
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(116592)
; OTHER INFORMATION: n = A,T,C or G
US-09-818-512-3

Query Match          52.7%; Score 21.6; DB 10; Length 116592;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 24; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CTTTGTGGTGATTTCTGCRGTGTGTAA 30
||||| ||| ||||| ||||| |||||
Db 23438 CTTTGTATGCCATCTTTCTTCATGTAA 23467

RESULT 9
US-09-864-761-1146/c
; Sequence 1146, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

Matches	28;	Conservative	1;	Mismatches	12;	Indels	0;	Gaps	0;
QY	1	CTTTGTGTGGGTGATTTCGTCGTCGTGTAATCGTGTCCCTG	41						
Db	2091	CTTTGTGTGTCGACTGTCGTGTGTGTGTGTGTGTGTGTGTGTG	2131						
RESULT	11								
US-09-799-462A-16									
;	Sequence 16, Application US/09799462A								
;	Patent No. US20020160970A1								
;	GENERAL INFORMATION:								
;	APPLICANT: Hadlaczky, Gyula								
;	Attorney: Szalay, Aladar								
;	TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES								
;	NUMBER OF SEQUENCES: 34								
;	CORRESPONDENCE ADDRESS:								
;	ADDRESSEE: Heller Ehrman White & McAuliffe								
;	STREET: 4250 Executive Square, 7th Floor								
;	CITY: La Jolla								
;	STATE: CA								
;	COUNTRY: USA								
;	ZIP: 92037								
;	COMPUTER READABLE FORM:								
;	MEDIUM TYPE: Diskette								
;	COMPUTER: IBM Compatible								
;	OPERATING SYSTEM: DOS								
;	SOFTWARE: FastSeq Version 1.5								
;	CURRENT APPLICATION DATA:								
;	APPLICATION NUMBER: US/09/799,462A								
;	FILING DATE: 10-Sep-2001								
;	CLASSIFICATION: <Unknown>								
;	PRIOR APPLICATION DATA:								
;	APPLICATION NUMBER: 08/835,682								
;	FILING DATE: 10-APR-1997								
;	APPLICATION NUMBER: 08/695,191								
;	FILING DATE: 07-AUG-1996								
;	APPLICATION NUMBER: 08/682,080								
;	FILING DATE: 15-JUL-1996								
;	APPLICATION NUMBER: 08/829,822								
;	FILING DATE: 10-APR-1996								
;	ATTORNEY/AGENT INFORMATION:								
;	NAME: Seidman, Stephanie L								
;	REGISTRATION NUMBER: 33,779								
;	REFERENCE/DOCKET NUMBER: 24601-402G								
;	TELECOMMUNICATION INFORMATION:								
;	TELEPHONE: 858-450-8403								
;	TELEFAX: 858-587-5360								
;	TELEX: <Unknown>								
;	INFORMATION FOR SEQ ID NO: 16:								
;	SEQUENCE CHARACTERISTICS:								
;	LENGTH: 2218 base pairs								
;	TYPE: nucleic acid								
;	STRANDEDNESS: single								
;	TOPOLOGY: linear								
;	MOLECULE TYPE: Genomic DNA								
;	HYPOTHETICAL: NO								
;	ANTI-SENSE: NO								
;	FRAGMENT TYPE: <Unknown>								
;	ORIGINAL SOURCE:								
;	SEQUENCE DESCRIPTION: SEQ ID NO: 16:								
US-09-799-462A-16									
Query Match	52.2%;	Score 21.4;	DB 9;	Length 22118;					
Best Local Similarity	68.3%;	Pred No. 61;							
Matches	28;	Conservative	1;	Mismatches	12;	Indels	0;	Gaps	0;
Qy	1	CTTTGTGTGGGTGATTTCGTCGTCGTGTAATCGTGTCCCTG	41						
bb	20136	CTTTG	20176						

```
RESULT 12
US-10-125-767-16
; Sequence 16, Application US/10125767
; Patent No. US20020160410A1
; GENERAL INFORMATION:
; APPLICANT: Hadlaczky, Gyula
; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND
; METHODS
; FOR PREPARING ARTIFICIAL CHROMOSOMES
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe LLP
; STREET: 4350 La Jolla Village Drive, 7th Floor
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/125,767
; FILING DATE: 17-Apr-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/724,693
; FILING DATE: 28-NOV-2000
; APPLICATION NUMBER: 08/835,682
; FILING DATE: 10-APR-1997
; APPLICATION NUMBER: 08/695,191
; FILING DATE: 07-AUG-1996
; APPLICATION NUMBER: 08/682,080
; FILING DATE: 15-JUL-1996
; APPLICATION NUMBER: 08/629,822
; FILING DATE: 10-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 24601-402J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858-450-8403
; TELEFAX: 858-587-5360
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22118 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: <Unknown>
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-125-767-16
Query Match 52.2%; Score 21.4; DB 9; Length 22118;
Best Local Similarity 68.3%; Pred. No. 61;
Matches 28; Conservative 1; Mismatches 12; Indels 0; Gaps 0;
QY 1 CTTTGTGGTGGGATTTTCTGCTGTGTAATCGTGCCCTG 41
||||| ||| | | | | | | | | | | | | | | | | |
Db 20136 CTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 20176

RESULT 14
US-09-771-208-20/c
; Sequence 20, Application US/09771208
; Patent No. US20020155564A1
; GENERAL INFORMATION:
; APPLICANT: MEDRANO, JUAN
; APPLICANT: BRADFORD, ERIC
; APPLICANT: HORVAT, SIMON
; TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
; FILE REFERENCE: 407T-923710US
; CURRENT APPLICATION NUMBER: US/09/771,208
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 08/999,477
; PRIOR FILING DATE: 1997-12-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (123459)..(123478)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (602466)..(602485)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (546998)..(547017)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (494715)..(494814)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (390986)..(391005)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (346860)..(346823)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (317174)..(317193)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (280353)..(280373)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (271829)..(271848)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (183872)..(183891)
; OTHER INFORMATION: n is unidentified a, c, g, or t
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:43:59 ; Search time 16.6596 Seconds  
(without alignments)  
5542.256 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_600\_640

Perfect score: 41  
Sequence: 1 agaaagcagtggaggagac.....accctcaggcagcccgaggag 41

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

1:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24:	/SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.6	99.0	1669	22	AAH26169 Human cytochrome P
2	40.6	99.0	1669	22	AAH26179 Human cytochrome P
3	40.6	99.0	1680	24	AAD34214 Human CYP2D6 gene
4	40.6	99.0	9432	24	AAD34213 Human cytochrome P
5	24.6	60.0	159	22	ABA75974 Human foetal liver
6	24.6	60.0	159	22	ABA75974 Human foetal liver
7	24.6	60.0	159	22	ABA40537 Probe #19003 for g
8	24.6	60.0	159	22	AAK24652 Human brain expres
9	24.6	60.0	159	22	AAK50649 Human bone marrow
					Probe #17595 for g

10	24.6	60.0	159	22	AAI56630 Probe #25316 used
11	24.6	60.0	159	24	ABS24124 Human genome-deriv
12	24.6	60.0	555	22	ABA63649 Human foetal liver
13	24.6	60.0	555	22	ABA30842 Probe #9308 for qe
14	24.6	60.0	555	22	AAK12168 Human brain expres
15	24.6	60.0	555	22	AAK37888 Human bone marrow
16	24.6	60.0	555	22	AAI18650 Probe #8583 for qe
17	24.6	60.0	555	22	AAI43763 Probe #12449 used
18	24.6	60.0	555	24	ABS11885 Human genome-deriv
19	24.6	60.0	2380	22	AAI60518 Human polynucleoti
20	24.6	60.0	2555	22	AAF59625 Human cell cycle a
21	24.6	60.0	2536	22	AAH08044 Human leucine-rich
22	24.6	60.0	2636	22	AAF90331 Nucleotide sequenc
23	24.6	60.0	2663	24	ABL89857 Human polynucleoti
24	24.6	60.0	2781	22	AAI58732 Human polynucleoti
25	24.6	60.0	3071	24	ABA93718 Human transmembran
26	23.6	57.6	121	16	AAT24091 Human gene signatu
27	23	56.1	2770	23	ABL18291 Drosophila melanog
28	23	56.1	9457	23	ABL18290 Drosophila melanog
29	22.4	54.6	2875	22	AAI41903 Genomic sequence #
30	22.4	54.6	3250	23	AA584219 Human cDNA encodin
31	22.4	54.6	3314	22	AA526311 Human cDNA encodin
32	22.4	54.6	4097	22	AA541904 Genomic sequence #
33	22	53.7	6868	24	ABK52418 DNA encoding propi
34	22	53.7	4403765	22	AAI99683 Mycobacterium tube
35	22	53.7	4411529	22	AAI99682 Mycobacterium tube
36	21.8	53.2	195	22	ABA47874 Human breast cell
37	21.8	53.2	195	22	ABA48709 Human breast cell
38	21.8	53.2	195	22	ABA65766 Human foetal liver
39	21.8	53.2	195	22	ABA66622 Human foetal liver
40	21.8	53.2	195	22	ABA32853 Probe #11319 for g
41	21.8	53.2	195	22	ABA33687 Probe #12153 for g
42	21.8	53.2	195	22	AAK14168 Human brain expres
43	21.8	53.2	195	22	AAK15049 Human brain expres
44	21.8	53.2	195	22	AAK39504 Human bone marrow
45	21.8	53.2	195	22	AAK40779 Human bone marrow

ALIGNMENTS

RESULT 1  
AAH26169  
ID AAH26169 standard; DNA; 1669 BP.  
XX  
AC AAH26169;  
XX  
DT 17-SEP-2001 (first entry)  
XX  
DE Human cytochrome P450 CYP2D6 gene promoter region.  
XX  
KW Cytochrome P450; CYP2D6; promoter; drug metabolism; human;  
diagnosis; therapy; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT primer\_bind Location/Qualifiers  
FT complement (14...36)  
FT /tag= a  
FT /note= "amplification primer upfl4"  
FT primer\_bind 337...358  
FT /tag= b  
FT /note= "sequence primer R1"  
FT primer\_bind 493...514  
FT /tag= c  
FT /note= "sequencing primer R2"  
FT primer\_bind complement (565..577)  
FT /tag= d  
FT /note= "sequencing primer F2"  
FT primer\_bind 602..620  
FT /tag= e  
FT /note= "sequencing primer R3"  
FT primer\_bind complement (968..988)





CC different alleles in human individuals, provide a very potent tool  
CC for improving the therapy of diseases with drugs that are targets  
CC of the CYP2D6 gene production, and whose metabolism is therefore  
CC dependent on CYP2D6 activity.  
CC Note: The present sequence is not shown in the specification but is  
CC derived from the CYP2D6 promoter sequence given in the Sequence  
CC Listing (see AH26169).  
XX  
SQ Sequence 1669 BP; 413 A; 376 C; 535 G; 338 T; 7 other;  
Query Match 99.0%; Score 40.6; DB 22; Length 1669;  
Best Local Similarity 100.0%; Pred. No. 6.4e-05;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGGCAGCCCGGGAG 41  
|||||  
DB 600 AGAAGCAGTGGAGGAGGACRACCCCTCAGGCAGCCCGGGAG 640  
RESULT 3  
AAD34214  
ID AAD34214 standard; DNA; 1680 BP.  
AC AAD34214;  
XX  
DT 16-JUL-2002 (first entry)  
XX  
DE Human CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX  
KW Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; polymorphism; ds.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT misc\_feature 36  
FT /\*tag= a  
FT /note= "Polymorphic site"  
FT misc\_feature 194  
FT /\*tag= b  
FT /note= "Polymorphic site"  
FT misc\_feature 385  
FT /\*tag= c  
FT /note= "Polymorphic site"  
FT misc\_feature 620  
FT /\*tag= d  
FT /note= "Polymorphic site"  
FT misc\_feature 880  
FT /\*tag= e  
FT /note= "Polymorphic site"  
FT misc\_feature 942  
FT /\*tag= f  
FT /note= "Polymorphic site"  
FT misc\_feature 1255  
FT /\*tag= g  
FT /note= "Polymorphic site"  
XX  
PN WO200218638-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 27-AUG-2001; 2001WO-IB01544.  
XX  
PR 30-AUG-2000; 2000GB-0021286.  
XX  
PA (GEMI-) GEMINI GENOMICS PLC.  
XX  
PI Risinger C, Andersson MK, Lewander T, Olliasson E;  
XX  
DR WPI; 2002-329785/36.  
XX  
PT New sequence determination oligonucleotides, useful for detecting

PT polymorphic sites in a 5' flanking region of a CYP2D6 gene, as  
PT hybridization probes, as components of diagnostic assays, or in  
PT ligase-based sequence determination -  
PS  
XX Claim 1; Fig 2; 63pp; English.  
XX  
CC The invention relates to sequence determination oligonucleotides for  
CC detecting polymorphic sites in a 5' flanking region of cytochrome P450  
CC 2D6 (CYP2D6) gene. CYP2D6 enzymes are involved in the metabolism of many  
CC different xenobiotics. Human CYP2D6 gene is located on chromosome 22.  
CC The oligonucleotides may be used as in situ hybridisation probes, in  
CC ligase-based sequenced determination, as components of diagnostic assays,  
CC as probes in sequence determination methods based on mismatches, as  
CC hybridisation-based diagnostic assays, and as components of diagnostic  
CC microarray. CYP2D6 is useful to predict variations in an individual's  
CC ability to metabolise certain drugs. The present sequence is human  
CC CYP2D6 gene 5' flanking region containing polymorphic sites.  
XX  
SQ Sequence 1680 BP; 413 A; 379 C; 539 G; 342 T; 7 other;  
Query Match 99.0%; Score 40.6; DB 24; Length 1680;  
Best Local Similarity 100.0%; Pred. No. 6.4e-05;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGGCAGCCCGGGAG 41  
|||||  
DB 600 AGAAGCAGTGGAGGAGGACRACCCCTCAGGCAGCCCGGGAG 640  
RESULT 4  
AAD34213  
ID AAD34213 standard; DNA; 9432 BP.  
XX  
AC AAD34213;  
XX  
DT 16-JUL-2002 (first entry)  
XX  
DE Human cytochrome P450 2D6 (CYP2D6) gene.  
XX  
KW Human; cytochrome P450 2D6; CYP2D6; enzyme; detection; xenobiotic;  
KW ligase-based sequenced determination; drug metabolism; chromosome 22;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200218638-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 27-AUG-2001; 2001WO-IB01544.  
XX  
PR 30-AUG-2000; 2000GB-0021286.  
XX  
PA (GEMI-) GEMINI GENOMICS PLC.  
XX  
PI Risinger C, Andersson MK, Lewander T, Olliasson E;  
XX  
DR WPI; 2002-329785/36.  
XX  
PT New sequence determination oligonucleotides, useful for detecting

CC hybridisation-based diagnostic assays, and as components of diagnostic  
CC microarray. CYP2D6 is useful to predict variations in an individual's  
CC ability to metabolise certain drugs. The present sequence is human  
CC CYP2D6 gene.  
XX  
SQ Sequence 9432 BP; 1964 A; 2647 C; 2976 G; 1845 T; 0 other;  
  
Query Match 99.0%; Score 40.6; DB 24; Length 9432;  
Best Local Similarity 97.6%; Pred. No. 7.2e-05;  
Matches 40; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AGAAAGCAGTGGAGGAGGACRACCCCTCAGGACGCCGGGAG 41  
|||||  
Db 600 AGAAAGCAGTGGAGGAGGACGACCCCTCAGGACGCCGGGAG 640  
|||||  
  
RESULT 5  
ABA75974  
ID ID ABA75974 standard; DNA; 159 BP.  
XX  
AC ABA75974;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
XX Human foetal liver single exon nucleic acid probe #24279.  
DE Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200157277-A2.  
PN  
PD 09-AUG-2001.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PF Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483447/52.  
PR  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA  
PI  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-483447/52.  
DR  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -  
PT  
PS Claim 4; SEQ ID NO 24279; 639pp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 159 BP; 38 A; 50 C; 51 G; 20 T; 0 other;  
  
Query Match 60.0%; Score 24.6; DB 22; Length 159;  
Best Local Similarity 73.2%; Pred. No. 19;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 AGAAAGCAGTGGAGGAGGACRACCCCTCAGGACGCCGGGAG 41  
|||||  
Db 600 AGAAAGCAGTGGAGGAGGACGACCCCTCAGGACGCCGGGAG 640  
|||||

Db 49 AGAAGCCCTGCTGGAGGAGGACCCCTCAGGGTCCAGGAAG 89  
  
RESULT 6  
ABA40537  
ID ID ABA40537 standard; DNA; 159 BP.  
XX  
AC ABA40537;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
XX Probe #19003 for gene expression analysis in human heart cell sample.  
DE Human; gene expression; heart; microarray; vascular system; probe;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200157274-A2.  
PN  
PD 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00666.  
PF  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488899/53.  
DR  
XX Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts -  
PT  
XX Claim 4; SEQ ID No 19003; 530pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart. The  
CC present sequence is one such probe. The probes may be used for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from the human heart via microarrays. By measuring gene expression, the  
CC probes are useful for predicting, diagnosing, grading, staging,  
CC monitoring and prognosing diseases of the human heart and vascular system  
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
CC congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 159 BP; 38 A; 50 C; 51 G; 20 T; 0 other;  
  
Query Match 60.0%; Score 24.6; DB 22; Length 159;  
Best Local Similarity 73.2%; Pred. No. 19;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
  
QY 1 AGAAAGCAGTGGAGGAGGACRACCCCTCAGGACGCCGGGAG 41  
|||||  
Db 49 AGAAGCCCTGCTGGAGGAGGACCCCTCAGGGTCCAGGAAG 89  
|||||

```
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe SEQ ID NO: 24643.
XX PR Human: brain expressed exon; gene expression analysis; probe:
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00667.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PD WPI; 2001-483446/52.
XX PF Single exon nucleic acid probes for analyzing gene expression in human
XX PR brains -
XX PS Example 4; SEQ ID NO: 24643; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC epilepsy and cancers. The present sequence is one of the probes of the
XX CC invention.
XX SQ Sequence 159 BP; 38 A; 50 C; 51 G; 20 T; 0 other;
XX
XX Query Match 60.0%; Score 24.6; DB 22; Length 159;
XX Best Local Similarity 73.2%; Pred. No. 19;
XX Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
XX
XX QY 1 AGAAGCAGCTGGAGGAGGACRACCCCTCAGGCGCGGGAG 41
XX Db |||| | || |||||:||||||| ||| || ||
XX 49 AGAAGCCCTGCTGGAGGAGCGACCCCTCAGGTGCCAGGAAG 89
XX
XX RESULT 8
XX AAK50649
XX ID AAK50649 standard; DNA; 159 BP.
XX AC AAK50649;
XX XX
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 25206.
XX PR Human: bone marrow expressed exon; gene expression analysis; probe:
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX KW
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00670.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PD WPI; 2001-483446/52.
XX PF Single exon nucleic acid probes for analyzing gene expression in human
XX PR brains -
XX PS Example 4; SEQ ID NO: 24643; 650pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC epilepsy and cancers. The present sequence is one of the probes of the
XX CC invention.
XX SQ Sequence 159 BP; 38 A; 50 C; 51 G; 20 T; 0 other;
XX
XX Query Match 60.0%; Score 24.6; DB 22; Length 159;
XX Best Local Similarity 73.2%; Pred. No. 19;
XX Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
XX
XX QY 1 AGAAGCAGCTGGAGGAGGACRACCCCTCAGGCGCGGGAG 41
XX Db |||| | || |||||:||||||| ||| || ||
XX 49 AGAAGCCCTGCTGGAGGAGCGACCCCTCAGGTGCCAGGAAG 89
XX
XX RESULT 9
XX AAI27662
XX ID AAI27662 standard; DNA; 159 BP.
XX AC AAI27662;
XX XX
XX DT 12-OCT-2001 (first entry)
XX DE Probe #17595 for gene expression analysis in human cervical cell sample.
XX PR Probe: human; microarray; gene expression; cervical epithelial cell;
XX KW cervical cancer; ss.
XX OS Homo sapiens.
XX PN WO200157278-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00670.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PD WPI; 2001-488900/53.
XX PF Human genome-derived single exon nucleic acid probes useful for
XX PR analyzing gene expression in human bone marrow -
XX PS Example 4; SEQ ID NO: 25206; 658pp + Sequence Listing; English.
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX CC the probes of the invention.
XX SQ Sequence 159 BP; 38 A; 50 C; 51 G; 20 T; 0 other;
```

DR WPI; 2001-488901/53.  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
PS Claim 25; SEQ ID No 17595; 487pp; English.  
XX The present invention relates to human single exon nucleic acid probes  
CC (SENP). The present sequence is one such probe. The SENPs are derived  
CC from human HeLa cells. The SENPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging  
CC of diseases of the cervix, notably cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 159 BP; 38 A; 50 C; 51 G; 20 T; 0 other;  
Query Match 60.0%; Score 24.6; DB 22; Length 159;  
Best Local Similarity 73.2%; Pred. No. 19;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 1 AGAAAGCAGTGGAGGAGGACACCTCAGGCAGCCCGGGAG 41  
Db 49 AGAAGCCCTGCTGGAGGACGACCTCAGGCTGCCAGGAAG 89  
RESULT 10  
AAI56630  
ID AAI56630 standard; DNA; 159 BP.  
XX  
AC AAI56630;  
XX  
XX 17-OCT-2001 (first entry)  
XX  
DE Probe #25316 used to measure gene expression in human placenta sample.  
XX  
KW Probe; microarray; human; placenta; antenatal diagnosis;  
XX genetic disorder; ss.  
OS Homo sapiens.  
XX  
XX WO200157272-A2.  
XX  
XX 09-AUG-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00663.  
XX  
XX 04-FEB-2000; 2000US-0180312.  
XX 26-MAY-2000; 2000US-0207456.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-0234687.  
XX 27-SEP-2000; 2000US-0236359.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488897/53.  
XX Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
XX Claim 25; SEQ ID No 25316; 654pp; English.  
XX  
XX The present invention relates to single exon nucleic acid probes (SENP).  
CC The present sequence is one such probe. The probes are useful for  
CC producing a microarray for predicting, measuring and displaying gene  
CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders.  
XX  
SQ Sequence 159 BP; 38 A; 50 C; 51 G; 20 T; 0 other;  
Query Match 60.0%; Score 24.6; DB 22; Length 159;  
Best Local Similarity 73.2%; Pred. No. 19;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 1 AGAAAGCAGTGGAGGAGGACACCTCAGGCAGCCCGGGAG 41  
Db 49 AGAAGCCCTGCTGGAGGACGACCTCAGGCTGCCAGGAAG 89  
RESULT 11  
ABS24124  
ID ABS24124 standard; DNA; 159 BP.  
XX  
AC ABS24124;  
XX  
XX 19-AUG-2002 (first entry)  
XX Human genome-derived single exon probe ORF from lung SEQ ID No 24115.  
XX  
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
KW chronic obstructive pulmonary disease; interstitial lung disease;  
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenar syndrome;  
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease; open reading frame; ORF.  
XX  
OS Homo sapiens.  
XX  
XX WO200186003-A2.  
XX  
XX 15-NOV-2001.  
XX  
XX 30-JAN-2001; 2001WO-US00665.  
XX  
XX 04-FEB-2000; 2000US-180312P.  
XX 26-MAY-2000; 2000US-207456P.  
XX 30-JUN-2000; 2000US-0608408.  
XX 03-AUG-2000; 2000US-0632366.  
XX 21-SEP-2000; 2000US-234687P.  
XX 27-SEP-2000; 2000US-236359P.  
XX 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2002-114183/15.  
XX  
XX Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples -  
XX  
XX Claim 4; SEQ ID No 24115; 634pp; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of  
CC nucleic acid expressed in the human lung; measuring gene expression to a  
CC sample derived from human lung, comprising (a) contacting the array with  
CC mRNA, and (b) measuring the label detectably bound to each probe of  
CC the array; identifying exons in a eukaryotic genome, comprising  
CC (a) algorithmically predicting at least one exon from genomic sequences

PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human fetal liver -  
XX  
XX  
PS Claim 1; SEQ ID NO 11954; 639pp + sequence listing; English.  
XX  
XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at [ftp.wipo.int/pub/published\\_pct-sequences](http://ftp.wipo.int/pub/published_pct-sequences).  
XX  
XX Sequence 555 BP; 128 A; 158 C; 148 G; 121 T; 0 other;  
SQ

Query Match	60.08;	Score 24.6;	DB 22;	Length 555;
Best Local Similarity	73.2%;	Pred. No. 21;		
Matches 30;	Conservative	1;	Mismatches 10;	Indels 0;
				Caps 0;

Qy 1 AGAAGCAGTGGAGGAGGACRACCTCAGGCAGCCCCGGAG 41  
 |||| | || |||||:||||||| ||| ||| ||  
 Db 298 AGAAGCCCTGCTGGAGGAGACCCCTCAGGGTGCCAGGAG 338

RESULT 13  
ABA30842

ABA30842  
ID ABA30842 standard; DNA; 555 BP.

AC ABA30842;

DT 23-JAN-2002 (first entry)

Probe #9308 for gene expression analysis in human heart cell sample.

XX Human; gene expression; heart; microarray; vascular system; probe;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease; ss.

OS Homo sapiens.

PN WO200157274-A2.

XX  
09-AUG-2001XX  
DE 30-JAN-2001. 2001WO-JIS00666

XX  
01-SEP-2000. 2000US-0190313

PR 26-MAY-2000; 2000US-0207456.  
PR 30-MAY-2000; 2000US-0208408.  
PR 30-MAY-2000; 2000US-0208408.

PR 03-AUG-2000; 2000US-0632366.

PR 27-SEP-2000; 2000US-0236359.

XX  
XX  
04 001 2000, ZOOGBE-00Z7203.  
  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX

PI Penn SG, Hanzel DK,

XX  
WPB 3001-100000/53

DRK 2001-400099/33.  
XX

PT Single exon nucleic

PI hearts -  
XX

PS Claim 1; SEQ ID No 9

XX  
XX

CC The present invention  
CC measuring human gene

CC present sequence is

CC predicting, measuring  
CC from the human heart  
CC

CC probes are useful for



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 20:51:09 ; Search time 118.283 Seconds  
(without alignments)  
5613.764 Million cell updates/sec

Title: US-09-942-310-2\_COPY\_600\_640  
Perfect score: 41  
Sequence: 1 agaaagcagtgaggagac.....accctcagcgagccgggag 41

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfum:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	40.6	99.0	886	13	BI910844
2	34.2	83.4	1180	14	BQ067695
3	24.6	60.0	265	9	AA126901
4	24.6	60.0	334	14	BM796175
5	24.6	60.0	366	13	BG954027
6	24.6	60.0	379	12	BF755875

c 7	24.6	60.0	383	13	BG954034
8	24.6	60.0	402	9	AA351643
9	24.6	60.0	448	9	AI276126
c 10	24.6	60.0	464	13	BI033426
11	24.6	60.0	470	14	BM697806
12	24.6	60.0	506	14	BM674533
13	24.6	60.0	536	9	AI963082
14	24.6	60.0	540	14	BQ575427
15	24.6	60.0	541	13	BM040921
16	24.6	60.0	546	14	BM792240
17	24.6	60.0	554	12	BG222601
18	24.6	60.0	610	10	AW964936
c 19	24.6	60.0	620	12	BF345885
20	24.6	60.0	625	12	BF448000
21	24.6	60.0	630	14	BM728944
22	24.6	60.0	664	14	BM702091
c 23	24.6	60.0	666	17	AZ624786
24	24.6	60.0	688	12	BE789481
25	24.6	60.0	733	12	BG685391
26	24.6	60.0	752	13	BM040887
27	24.6	60.0	862	13	BM041161
28	24.6	60.0	902	14	BQ438013
29	24.6	60.0	914	12	BG400812
c 30	24.6	60.0	915	14	BQ718652
c 31	24.6	60.0	924	12	BF346206
32	24.6	60.0	1057	12	BG396658
c 33	24.6	60.0	1734	11	BC019839
34	24.4	59.5	652	10	AV927751
35	24.2	59.0	194	9	AA648675
c 36	24.2	59.0	1059	13	BM560193
37	23.2	56.6	393	17	AZ582757
38	23.2	56.6	402	17	AZ039011
c 39	23.2	56.6	445	10	AW185327
40	23.2	56.6	603	17	AZ703145
c 41	23.2	56.6	645	10	BE346825
c 42	23.2	56.6	699	10	BE346227
43	23.2	56.6	700	17	AZ808318
c 44	23	56.1	131	10	AW802995
c 45	23	56.1	199	10	AW434458

ALIGNMENTS

RESULT 1  
BI910844/c  
LOCUS BI910844 886 bp mRNA linear EST 16-OCT-2001  
DEFINITION 603068871F1 NIH\_MGC\_118 Homo sapiens cDNA clone IMAGE:5217977 5', mRNA sequence.  
ACCESSION BI910844  
VERSION BI910844.1 GI:16174292  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 886)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Plate: LLAM11547 row: m column: 18  
High quality sequence start: 4  
High quality sequence stop: 432.  
Location/Qualifiers







Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-CT0570-  
301000-432-h08&t3=2000-10-30&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 13  
High quality sequence stop: 379.

#### FEATURES

source  
1. .379  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CT0570"  
/dev\_stage="Adult"  
/note="Organ: colon; Vector: puc18; Site.1: SmaI; Site.2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."  
71 a 121 c 110 g 77 t

BASE COUNT 71 a 121 c 110 g 77 t  
ORIGIN  
Query Match 60.0%; Score 24.6; DB 12; Length 379;  
Best Local Similarity 73.2%; Pred. No. 2.4e+02;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
QY 1 AGAAGCAGTGAGGAGGACRACCCCTCAGCAGCCGGGAG 41  
||||| | | | | | | | | | | | | | | | | | | | | |  
Db 33 AGAAGCCCTGCTGGAGGAGCAGCCCTCAGGTCGCCAGGAAG 73

RESULT 7  
BG954034/c 383 bp mRNA linear EST 12-JUN-2001  
LOCUS  
DEFINITION CM3-CT0570-050201-716-h08 CT0570 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG954034  
VERSION BG954034.1 GI:14372205  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 383)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

TITLE  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM3&t2=CM3-CT0570-  
050201-716-h08&t3=2001-02-05&t4=1)

Seq primer: puc 18 forward  
High quality sequence stop: 383.  
FEATURES  
Location/Qualifiers  
source  
1. .383  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="CT0570"  
/dev\_stage="Adult"  
/note="Organ: colon; Vector: puc18; Site.1: SmaI; Site.2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No. 196  
716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."  
74 a 111 c 126 g 72 t

BASE COUNT 74 a 111 c 126 g 72 t  
ORIGIN

Query Match 60.0%; Score 24.6; DB 13; Length 383;  
Best Local Similarity 73.2%; Pred. No. 2.4e+02;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGAAACAGTGAGGAGGACRACCCCTCAGCAGCCGGGAG 41  
||||| | | | | | | | | | | | | | | | | | | | | |  
Db 345 AGAAGCCCTGCTGGAGGAGCAGCCCTCAGGTCGCCAGGAAG 305

RESULT 8  
AA351643 402 bp mRNA linear EST 21-APR-1997  
LOCUS  
DEFINITION EST59477 Infant brain Homo sapiens cDNA 5' end, mRNA sequence.  
ACCESSION AA351643  
VERSION AA351643.1 GI:2003983  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 402)  
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult  
C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White  
O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A.,  
Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald  
L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,  
Gnemm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr., Kelley,J.M.,  
Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,  
Moreno-Palanco,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,  
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,  
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,  
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,  
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,  
Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,  
Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,  
Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillon,P.J., Fannon  
M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and  
Venter,J.C.  
Initial assessment of human gene diversity and expression patterns  
based upon 83 million nucleotides of cDNA sequence

TITLE  
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)  
MEDLINE 96026280  
COMMENT Other\_ESTs: EST59476  
Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@tigr.org

For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (http://www.tigr.org/tdb/hgi/hgi.html)  
Seq primer: M13 Reverse.  
Location/Qualifiers

#### FEATURES

```
source
1. .402
/organism="Homo sapiens"
/db_xref="ATCC (inhost):151920"
/db_xref="taxon:9606"
/clone_lib="Infant brain"
/sex="female"
/dev_stage="infant"
/Note="Organ: Brain; Vector: lafmid BA; Site_1: HindIII; Site_2: NotI"
BASE COUNT      106 a 119 c 114 g 60 t 3 others
ORIGIN

Query Match      60.0%; Score 24.6; DB 9; Length 402;
Best Local Similarity 73.2%; Pred. No. 2.4e+02;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGGAGCCCGGGAG 41
||||| 1 || |||||:||||| 111 |||
Db 203 AGAAGCCCTGCTGGAGGAGCCCTCAGGTGCCAGGAAG 243
||||| 1 || |||||:||||| 111 |||

RESULT 9
AI276126      448 bp mRNA linear EST 29-JAN-1999
LOCUS      q170b05.x1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1877649
DEFINITION      3', mRNA sequence.
ACCESSION      AI276126
VERSION      AI276126.1 GI:3898400
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 448)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1079 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 433.
Location/Qualifiers
1. .448
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:1877649"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/Note="Organ: mixed (see below); Vector: pT73D-Pac (pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NBHPU, and fetal heart NBHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT      85 a 143 c 134 g 86 t
ORIGIN

Query Match      60.0%; Score 24.6; DB 9; Length 448;
Best Local Similarity 73.2%; Pred. No. 2.5e+02;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGGAGCCCGGGAG 41
```

```
Db 75 AGAAGCCCTGCTGGAGGAGCCCTCAGGTGCCAGGAAG 115
||||| 1 || |||||:||||| 111 |||

RESULT 10
BI033426/c      464 bp mRNA linear EST 14-JUN-2001
LOCUS      PM1-NN1207-150201-023-a01 NN1207 Homo sapiens cDNA, mRNA sequence.
ACCESSION      BI033426
VERSION      BI033426.1 GI:14440052
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 464)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM1&t2=PM1-NN1207-150201-023-a01&t3=2001-02-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 9
High quality sequence stop: 464.
Location/Qualifiers
1. .464
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NN1207"
/dev_stage="Adult"
/Note="Organ: nervous_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT      83 a 120 c 146 g 115 t
ORIGIN

Query Match      60.0%; Score 24.6; DB 13; Length 464;
Best Local Similarity 73.2%; Pred. No. 2.5e+02;
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGAAGCAGTGGAGGAGGACRACCCCTCAGGAGCCCGGGAG 41
||||| 1 || |||||:||||| 111 |||
Db 324 AGAAGCCCTGCTGGAGGAGCCCTCAGGTGCCAGGAAG 284
||||| 1 || |||||:||||| 111 |||

RESULT 11
BM697806      470 bp mRNA linear EST 28-FEB-2002
LOCUS      UI-E-DX0-ago-a-04-0-UI.r1 UI-E-DX0 Homo sapiens cDNA clone
DEFINITION      UI-E-DX0-ago-a-04-0-UI 5', mRNA sequence.
ACCESSION      BM697806
VERSION      BM697806.1 GI:19011064
```

```

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 470)
NORMALIZATION Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES
source
1..470
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-DX0-ago-a-04-0-UI"
/tissue_lib="UI-E-DX0"
/tissue_type="fetal eyes"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-DX0 is a cDNA library containing the following
tissue(s): fetal eyes. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (gt)18 tail. The
sequence tag for this library is AGAATCAAGA. This library
was created for the program Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
```

BASE COUNT 117 a 142 c 136 g 75 t

ORIGIN

Query Match 60.0%; Score 24.6; DB 14; Length 470;  
Best Local Similarity 73.2%; Pred. No. 2.5e+02;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGAAACAGTGGAGGAGGACACCTCAGGAGCCCGGGAG 41  
||||| | || |||||:||||||| ||| |||

Db 254 AGAAGCCCTGCTGGAGGAGGACACCTCAGGTTGCCAGGAAG 294

RESULT 12  
BM674533 506 bp mRNA linear EST 27-FEB-2002

LOCUS UI-E-EJ0-ahl-o-05-0-UI.s1 UI-E-EJ0 Homo sapiens cDNA clone

DEFINITION UI-E-EJ0-ahl-o-05-0-UI 3', mRNA sequence.

ACCESSION BM674533

VERSION BM674533.1 GI:18984431

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE 1 (bases 1 to 506)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
1..506
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-E-EJ0-ahl-o-05-0-UI"
/clone_lib="UI-E-EJ0"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stage="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-EJ0 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (drr)18 tail. The
sequence tags for this library are: fetal eyes, AGAATCAAGA
; lens, CGAATTAGCGA; eye anterior segment, AATGCCGCAAT;
optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and
Macular, GTCC; RPE and Choroid, ACCrA. This library was
created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
```

BASE COUNT 93 a 150 c 150 g 113 t

ORIGIN

Query Match 60.0%; Score 24.6; DB 14; Length 506;  
Best Local Similarity 73.2%; Pred. No. 2.5e+02;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGAAAGCAGTGGAGGAGGACRACCTCAGGAGCCCGGGAG 41  
||||| | || |||||:||||||| ||| |||

Db 93 AGAAGCCCTGCTGGAGGAGGACACCTCAGGTTGCCAGGAAG 133

RESULT 13  
AI963082 536 bp mRNA linear EST 09-MAR-2000

LOCUS wt27a05.x1 NCI\_CGAP\_Ut1 Homo sapiens cDNA clone IMAGE:2508656 3'

DEFINITION similar to TR:070210 070210 CHONDROADHERIN. ; mRNA sequence.

ACCESSION AI963082

VERSION AI963082.1 GI:5755795

KEYWORDS EST.

SOURCE human.



/lab\_host="DH10B (phage-resistant)"  
/note="Organ: kidney; Vector: pOTB7; Site.1: XhoI; Site.2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

BASE COUNT 152 a 139 c 167 g 83 t  
ORIGIN

Query Match 60.0%; Score 24.6; DB 13; Length 541;  
Best Local Similarity 73.2%; Pred. No. 2.6e+02;  
Matches 30; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 AGAAGCAGTGGAGGAGGACCCCTCAGCGCCCGGGGAG 41  
||||||| ||||||| :| ||||| || |||  
Db 192 AGAAGCAGAGGAGGAGGTCGAGGCTCAGAGAGGCGAGGGG 232

Search completed: February 11, 2003, 03:31:14  
Job time : 122.283 secs